

Untersuchung des Einflusses der farblichen Gestaltung der Bedienoberfläche auf den Stellwerksbediener

Von der
Fakultät Architektur, Bauingenieurwesen und Umweltwissenschaften
der Technischen Universität Carolo-Wilhelmina
zu Braunschweig

zur Erlangung des Grades eines
Doktor-Ingenieurs (Dr.-Ing.)
genehmigte

Dissertation

von
David Matz
geboren am 21. Februar 1985
aus Merseburg (Saale)

Eingereicht am: 23. Juni 2017

Disputation am: 27. Oktober 2017

Berichterstatter: Prof. Dr. Jörn Pachl
Prof. Dr. Mark Vollrath

2017

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Architektur, Bauingenieurwesen und Umweltwissenschaften, vertreten durch den Betreuer der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Tagungsbeiträge

Matz, D.; Milius, B.: Measurements and Data Collecting for Assessing Usability and User Experience of Interlocking Interfaces. In Milius, B.; Naumann, A. (Hrsg.): *Rail Human Factors. Proceedings of the 2nd German Workshop on Rail Human Factors*. ITS Automotive Nord, Braunschweig (2016)

Fachzeitschriften

Matz, D.: Messung der User Experience am Fahrdienstleiterarbeitsplatz. In ZEVrail – Zeitschrift für das gesamte System Bahn 140 (2016), Nr. 6-7, S. 239–243

VORWORT UND DANKSAGUNG

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Eisenbahnwesen und Verkehrssicherung der Technischen Universität Braunschweig. An dieser Stelle möchte ich die Gelegenheit nutzen, den Menschen zu danken, die durch ihre Unterstützung zum Gelingen dieser Arbeit beigetragen haben. Allen voran danke ich meinem Doktorvater Prof. Dr.-Ing. Jörn Pachl für die Hilfestellung bei fachlichen Fragen, die kontinuierliche Begleitung und den nötigen Freiraum.

Mein Dank gilt ebenso Prof. Dr. Mark Vollrath, Leiter des Fachgebiets Ingenieur- und Verkehrspsychologie am Institut für Psychologie der Technischen Universität Braunschweig, für die Erstellung des Zweitgutachtens und für die Beratung zur statistischen Auswertung.

Prof. Dr.-Ing. Thomas Siefer, Leiter des Instituts für Verkehrswesen, Eisenbahnbau und -betrieb der Technischen Universität Braunschweig, danke ich für die Übernahme des Vorsitzes der Prüfungskommission.

Für die Erstellung der verwendeten Bedienoberflächen bedanke ich mich bei Marc Burkhardt von der Siemens AG. Ohne die technische Unterstützung wäre die Arbeit nicht realisierbar gewesen.

Bedanken möchte ich mich bei allen Kollegen am Institut für die angenehme Zusammenarbeit. Dabei danke ich vor allem meiner Bürokollegin und Freundin Dr. Anne Lorenz für die vielen inspirierenden Gespräche und die sehr guten Canapés. Ein großes Dankeschön auch an Dr. Birgit Milius für die hilfreichen Anmerkungen und die konstruktive Kritik.

Weiter möchte ich mich bei den studentischen Hilfskräften Phillipp Sell, Jonas Flemming und Dominik Knebel bedanken, die während meiner Zeit am Institut zu Freunden wurden. Ich bedanke mich für ihre tatkräftige Unterstützung. Ein ganz besonderer Dank gebührt dabei Dominik Knebel, der mich während der Versuchsdurchführung unterstützt und begleitet hat. Vielen Dank für die in die Auswertung der Daten investierte Zeit und die mehr als notwendige Ablenkung neben der Arbeit.

Meinem Freund Claus Scherbarth möchte ich für das offene Ohr danken und dass er mir stets durch Anregungen und Ratschläge beiseite stand, auch dann noch, als Teile der Arbeit sehr kurzfristig korrigiert werden mussten.

Großer Dank gebührt meinen Eltern für das Vertrauen und die grenzenlose Unterstützung in den letzten Jahren. Ohne euch wäre dies nie möglich gewesen.

Zu guter Letzt möchte ich mich bei meiner Lebenspartnerin Kerstin Schmitt für die unermüdliche Unterstützung von ganzem Herzen bedanken. Danke für die mit großer Sorgfalt durchgeführte Korrekturarbeit und den Motivationsschub in der Endphase.

KURZFASSUNG

Der Fahrdienstleiter ist für die sichere und planmäßige Durchführung des gesamten Bahnverkehrs innerhalb seines Zuständigkeitsbereiches verantwortlich und hat demnach eine zentrale Funktion im Eisenbahnwesen. Aufgrund zunehmender Automatisierung am Fahrdienstleiterarbeitsplatz ergeben sich neuartige Herausforderungen, welche Ausgangspunkt für die Überlegungen und Untersuchungen in dieser Arbeit sind.

Gerade in der Rückfallebene – also im Störfall – besteht durch den hohen Automatisierungsgrad die Gefahr von fahrdienstlichen Fehlhandlungen. Diese werden dadurch begünstigt, dass der Fahrdienstleiter im Regelbetrieb aus der aktiven Betriebsdurchführung ausgeschlossen wird, aber in der Rückfallebene ein integraler Bestandteil der Handlungskette ist. Durch diesen Ausschluss wird ihm die gesamthafte Erfassung der Betriebssituation erschwert. Aus diesem Grund ist es unerlässlich, seine Arbeitsmittel der neuen Situation entsprechend anzupassen.

Dass bei der Systementwicklung neben den funktionalen Aspekten auch die Betrachtung menschlicher Fähigkeiten und persönlicher Bedürfnisse von Interesse sind, wird mit der vorliegenden Arbeit gezeigt. Es wurde eine experimentelle Bedienoberflächenstudie ($N = 52$) durchgeführt, um den Einfluss der farblichen Gestaltung der Bedienoberfläche auf den Stellwerksbediener detaillierter beschreiben zu können. Hierfür wurden zwei Bedienoberflächen einer deutschen Stadtbahn miteinander verglichen. Diese unterschieden sich nur hinsichtlich der farblichen Gestaltung, nicht aber in Bezug auf Benutzbarkeit und Funktionalität. Die Hauptaufgabe der Versuchsteilnehmer bestand in der Überwachung des Bedienbereichs und in dem Erkennen von Unregelmäßigkeiten.

Die gefundenen Unterschiede ließen sich zum größten Teil auf das hohe Kontrastverhältnis der jeweiligen Farbkombinationen und die damit verbundene gute Bewertung der Attraktivität und der pragmatischen Qualität der Bedienoberfläche zurückführen. Zudem zeigte sich ein Zusammenhang zwischen der Ästhetik der Bedienoberfläche und der Bewertung der Usability. Eine gute Usability hatte Auswirkungen auf die Leistungsbereitschaft und das Überwachungsverhalten der Versuchsteilnehmer. Hierbei spielte vor allem die Durchschaubarkeit der Bedienoberfläche eine entscheidende Rolle. Des Weiteren wurden geschlechterspezifische Unterschiede festgestellt, welche ebenfalls dafür sprechen, dass moderne Bedienoberflächen auf die Bedürfnisse der Bediener angepasst werden müssen und persönliche Vorlieben bei der Entwicklung einbezogen werden sollten.

Die Erkenntnisse der Arbeit sind aus wissenschaftlicher Sicht relevant und für die Praxis von Bedeutung. Sie liefern einen Beitrag zur Entwicklung zukünftiger Bedienoberflächen und bilden die Basis für ein besseres Verständnis der Zusammenhänge sowie für weitere Forschungsbestrebungen.

ABSTRACT

The operator is responsible for the safe and punctual railway operation within his/her area of responsibility. Therefore, the operator is an important part of the railway sector. Due to increasing automation at the operator's workplace new challenges have evolved, which need to be considered.

During normal operations, control of the trains is ensured by the technical systems. The task of the operator is limited to monitor the systems. However, once the technical systems are not available or cannot be used due to, e.g. restrictions, the operator becomes an integral part of train operations and his actions might be immediately safety relevant. This can lead to an overall higher risk of human error as due to the missing integration into normal operations problems regarding, e.g. situation awareness cannot arise. For this reason, it is essential to adapt the operator's workplace as well as operational processes to the new situation.

An experimental user interface study was conducted in order to describe the influence of the colour design of the user interface on the operator. For this purpose, two user interfaces based on an existing interface of a German urban railway were compared at the virtual railway laboratory at the Institute of Railway Systems Engineering and Traffic Safety. The only difference between both of them was their colour design. Participants of the experiment were divided into two groups, each group using one of the interfaces. All participants had to monitor and control the area displayed on their screens and had to detect malfunctions of the system. As a result of the experiment, it could be shown that the performance as well as the impact of the tasks on the participants differed between both groups.

The perspicuity of the user interface played a decisive role. Differences in the results could be attributed to the high contrast ratio of the respective colour combinations. This also led to a good evaluation of the attractiveness and the pragmatic quality (usability) of the user interface. A dependency between the perceived aesthetics of the user interface and the usability assessment could be proven. This in turn led to a better motivation and had a positive impact on the monitoring behaviour of the participants. Furthermore, gender specific differences were identified, which indicates that modern user interfaces must be adapted to the needs of the user and should allow to take personal preferences into account.

The results of this thesis are relevant to the research area as well as for practicable application. They can contribute to the development of future user interfaces and form the basis for a better understanding of the dependencies between the functional and non-functional aspects.

„[...] Einer da vorne am bebenden Hebel,
eisern für alle das Schicksal bewacht!“

Der Lokomotivführer
Das glitzernde Band, Nr. 7, 1937

INHALTSVERZEICHNIS

Inhaltsverzeichnis	I
Abkürzungsverzeichnis.....	V
Abbildungsverzeichnis	VII
Tabellenverzeichnis	XI
1 Einleitung.....	1
1.1 Motivation	1
1.2 Zielsetzung.....	2
1.3 Vorgehensweise.....	3
2 Vorbetrachtungen und Hintergründe.....	5
2.1 Mensch und Technik.....	5
2.1.1 Einführung	5
2.1.2 Menschliche Fehler	12
2.1.3 Exkurs Human Factors im Eisenbahnwesen.....	18
2.1.4 Zusammenfassung Kapitel 2.1	22
2.2 Ansätze und Methoden zur Bewertung menschengerechter Gestaltung	23
2.2.1 Ansätze der menschengerechten Gestaltung.....	24
2.2.2 Ansätze zur Bewertung der Benutzerfreundlichkeit.....	29
2.2.3 Zusammenfassung Kapitel 2.2	35
2.3 Einfluss von Farbe und Kontrast.....	36
2.3.1 Vorstellung der Attribute Farbe und Kontrast	36
2.3.2 Vorgaben für barrierefreie Informationstechnik	38
2.3.3 Zusammenfassung Kapitel 2.3	42
2.4 Die Darstellungsarten	42
2.4.1 Gegenüberstellung	43
2.4.2 Einführung der Polarität.....	44
2.4.3 Aktuelle Situation bei der DB Netz AG	45
2.4.4 Zusammenfassung Kapitel 2.4	46

3	Zielsetzung der Arbeit	47
3.1	Einfluss der farblichen Gestaltung.....	48
3.1.1	User Experience.....	48
3.1.2	Psychische Beanspruchung	48
3.1.3	Befindlichkeit.....	49
3.1.4	Bedienungssicherheit	49
3.1.5	Fehlerentdeckungszeit	50
3.2	Operationalisierung der Bedienereffizienz.....	50
3.2.1	Übersicht der vordefinierten Leistungskennzahlen	51
3.2.2	Untersuchung der Mausbewegung.....	51
4	Die Versuchsumgebung	53
4.1	Allgemeine Beschreibung des Bedienplatzsystems	53
4.2	Die verwendeten Bedienplätze der Stadtbahn	54
4.2.1	Original- und Experimentalbedienplatz	55
4.2.2	Designbedienplatz	56
4.2.3	Erklärung der Ausleuchtung.....	57
4.2.4	Polarität und Barrierefreiheit der Bedienplätze	58
4.3	Betrieb in der Rückfallebene	62
4.3.1	Hilfshandlungen und Kommandofreigabe	62
4.3.2	Unregelmäßigkeiten während des Versuchs	63
4.3.3	Fehlerstatusanzeige	67
5	Bedienoberflächenstudie	69
5.1	Methode	69
5.1.1	Stichprobe	70
5.1.2	Versuchsplan	71
5.1.3	Material und Geräte.....	73
5.1.4	Operationalisierung der Bedienereffizienz	85
5.1.5	Durchführung (Versuchsablauf)	90
5.1.6	Statistische Auswertung.....	94

5.2	Ergebnisse der Bedienoberflächenstudie.....	96
5.2.1	Einfluss der farblichen Gestaltung	96
5.2.2	Einfluss des Geschlechts der Versuchsteilnehmer.....	119
5.2.3	Wechselwirkung zwischen den unabhängigen Variablen.....	135
5.3	Diskussion	146
5.3.1	Einfluss der farblichen Gestaltung	146
5.3.2	Einfluss des Geschlechts der Versuchsteilnehmer.....	153
5.3.3	Weiterführende Diskussion.....	154
5.3.4	Mögliche versuchsbedingte Einschränkungen.....	156
5.3.5	Zusammenfassung der Abhängigkeiten.....	157
6	Zusammenfassung	159
	Literaturverzeichnis	163
	Anhang.....	175

ABKÜRZUNGSVERZEICHNIS

AOI	Areas of Interest
BMS	Erfassungsbogen Belastung-Monotonie-Sättigung/Stress
CCOHS	Canadian Centre for Occupational Health and Safety
CUE-Model	Components of User Experience Model
DB AG	Deutsche Bahn Aktiengesellschaft
DB Netz AG	Deutsche Bahn Netz Aktiengesellschaft
DIN	Deutsches Institut für Normung e. V.
DSTW	Digitales Stellwerk
EN	Europäische Norm
ESTW	Elektronisches Stellwerk
ETCS	European Train Control System
EULYNX	European Initiative Linking Interlocking Subsystems
EZ-Skala	Eigenzustandsskala
Fdl	Fahrdienstleiter
FS	Fahrstraße
FU	Farbunterschied
GAHA	Gleisabschnitt Hilfsauflösen
GS	Gute-Schlechte-Stimmung (Dimension des MDBF)
HD	Helligkeitsdifferenz
HMI	Human-Machine-Interface
HQ	Hedonische Qualität
iBS	integriertes Bedienplatzsystem
IEA	International Ergonomics Association
IfEV	Institut für Eisenbahnwesen und Verkehrssicherung
IP	Ingenieurpsychologie
ISO	Internationale Organisation für Normung
KF1	Kommandofreigabe 1
KF2	Kommandofreigabe 2
KFV	Klick-Fehler-Verhältnis
KS	Konformitätsstufe nach WCAG 2.0
KV	Kontrastverhältnis
LSB	Leipziger Stimmungsbogen

LZB	Linienförmige Zugbeeinflussung
M	Mean (Mittelwert)
MDBF	Mehrdimensionale Befindlichkeitsfragebogen
MMI	Mensch-Maschine-Interaktion
MMS	Mensch-Maschine-System
MP	Messpunkt
NeuPro	Neuausrichtung Produktionssteuerung
OOTLUF	Out-Of-The-Loop-Unfamiliarity
PANAS	Positive and Negative Affectivity Scale
PQ	Pragmatische Qualität
RiM	Rail-in-Motion
RSSB	Rail Safety and Standards Board
RU	Ruhe-Unruhe (Dimension des MDBF)
SA	Situation Awareness (Situationsbewusstsein)
SD	Standard Deviation (Standardabweichung)
SES	Signal Entsperren
Sifa	Sicherheitsfahrschaltung
SOFI	Swedish Occupational Fatigue Inventory
SURT	Surrogate Reference Task
Tf	Triebfahrzeugführer
UX	User Experience
VDI	Verein Deutscher Ingenieure
V-EBL	Virtuelles Eisenbahnbetriebslabor
VICOS OC	Vehicle and Infrastructure Control and Operating System/Operation Management System Commuter
VICOS T&S	VICOS Training and Simulation
VS	Vorsichtssignal Stellen
W3C	World Wide Web Consortium
WCAG	Web Content Accessibility Guidelines
WHA	Weiche Hilfsauflösen
WHU	Weiche Hilfsumstellen
WM	Wachheit-Müdigkeit (Dimension des MDBF)
ZLB	Zuglenkbetrieb

ABBILDUNGSVERZEICHNIS

Abbildung 2-1: Darstellung der vereinfachten Systemstruktur eines ESTW	8
Abbildung 2-2: Die drei Ebenen des Situationsbewusstseins	10
Abbildung 2-3: Zusammenhang der Konzepte <i>complacency</i> und <i>automation bias</i>	12
Abbildung 2-4: Schematische Darstellung der kognitiven Ebenen des menschlichen Verhaltens	16
Abbildung 2-5: Einteilung sicherheitsgefährdender Handlungen	17
Abbildung 2-6: Basiswissenschaften der Human-Factors-Forschung	19
Abbildung 2-7: Das vereinfachte Strukturbild des Mensch-Maschine-Systems	25
Abbildung 2-8: Das erweiterte Strukturbild des Mensch-Maschine-Systems	25
Abbildung 2-9: Lage der Benutzerschnittstelle im modernen Stellwerk	27
Abbildung 2-10: Zusammenhang der drei UX-Qualitäten	33
Abbildung 2-11: CUE-Modell	34
Abbildung 2-12: Bestimmung der relativen Farbhelligkeit (Formel 2-1)	39
Abbildung 2-13: Bestimmung des Kontrasts (Formel 2-2)	39
Abbildung 2-14: Kontrast der Farben <i>Schwarz</i> und <i>Weiß</i>	39
Abbildung 2-15: Beispiel für Konformitätsstufe AA	40
Abbildung 2-16: Beispiel für Konformitätsstufe AA+	40
Abbildung 2-17: Bestimmung der Helligkeitsdifferenz (Formel 2-3).....	41
Abbildung 2-18: Bestimmung des Farbunterschieds (Formel 2-4).....	41
Abbildung 2-19: Positiv- vs. Negativdarstellung mit achromatischen Farben	43
Abbildung 4-1: Bedienoberfläche OC 111 (Originalbedienplatz)	55
Abbildung 4-2: Bedienoberfläche OC 111 (Experimentalbedienplatz)	55
Abbildung 4-3: Bedienoberfläche OC 111 (Designbedienplatz)	56
Abbildung 4-4: Darstellung weiterer Informationen über ein Pop-up-Fenster	57
Abbildung 4-5: Ausschnitt des verwendeten Lupenbilds (Originalbedienplatz).....	58
Abbildung 4-6: Dialogfenster für die Kommandofreigabe	63
Abbildung 4-7: Darstellung Ausfall Zuglenkbetrieb S27	64
Abbildung 4-8: Darstellung Signalstörung S27 (Signalkopf blinkt)	64
Abbildung 4-9: Darstellung Gleisfreimeldestörung W04	66
Abbildung 4-10: Darstellung Fehlerstatusanzeige.....	67
Abbildung 5-1: Aufteilung Teilnehmer auf die Gruppen und nach Geschlecht	70

Abbildung 5-2: Auszug aus dem UEQ.....	74
Abbildung 5-3: Darstellung Antwortformat (exemplarisch Frage 1).....	77
Abbildung 5-4: Auszug aus der EZ-Skala.....	79
Abbildung 5-5: Auszug aus dem MDBF (Kurzform A).....	80
Abbildung 5-6: Darstellung der Fixationsdauer, des Fixationsorts und der Blickpfade	82
Abbildung 5-7: Darstellung der Nebenaufgabe.....	84
Abbildung 5-8: Ausschnitt aus der JPEG-Datei, die durch das Tool <i>IOGraph</i> erzeugt wird.....	85
Abbildung 5-9: Konzept zur Operationalisierung der Bedienereffizienz.....	86
Abbildung 5-10: Bestimmung des Klick-Fehler-Verhältnisses (Formel 5-1).....	88
Abbildung 5-11: Mittelwerte der sechs Dimensionen bezogen auf die farbliche Gestaltung	97
Abbildung 5-12: Mittelwerte der UX-Qualitäten auf die farbliche Gestaltung.....	98
Abbildung 5-13: Mittelwerte der sieben Items bezogen auf die farbliche Gestaltung.....	99
Abbildung 5-14: Mittelwerte der Differenzen der Dimensionen bez. auf farbliche Gestaltung.....	105
Abbildung 5-15: Mittelwerte <i>Gute-Schlechte-Stimmung</i> bez. auf die farbliche Gestaltung.....	106
Abbildung 5-16: Mittelw. positiver Aspekt <i>Gute-Schlechte-Stimmung</i> bez. auf farbl. Gestaltung.....	106
Abbildung 5-17: Mittelw. negativer Aspekt <i>Gute-Schlechte-Stimmung</i> bez. auf farbl. Gestaltung....	107
Abbildung 5-18: Mittelwerte <i>Wachheit-Müdigkeit</i> bez. auf die farbliche Gestaltung.....	107
Abbildung 5-19: Mittelw. positiver Aspekt <i>Wachheit-Müdigkeit</i> bez. auf farbl. Gestaltung.....	108
Abbildung 5-20: Mittelw. negativer Aspekt <i>Wachheit-Müdigkeit</i> bez. auf farbl. Gestaltung.....	108
Abbildung 5-21: Mittelwerte <i>Ruhe-Unruhe</i> bez. auf die farbliche Gestaltung	109
Abbildung 5-22: Mittelw. positiver Aspekt <i>Ruhe-Unruhe</i> bez. auf farbl. Gestaltung.....	110
Abbildung 5-23: Mittelw. negativer Aspekt <i>Ruhe-Unruhe</i> bez. auf farbl. Gestaltung	110
Abbildung 5-24: Mittelwerte Messpunkte für Bedienungssicherheit bez. auf farbl. Gestaltung.....	111
Abbildung 5-25: Mittelwerte Messpunktdifferenzen bezogen auf die farbliche Gestaltung	112
Abbildung 5-26: Aufteilung Teilnehmer auf Gruppen und nach Geschlecht (reduzierte Größe)	113
Abbildung 5-27: Mittelwerte Fehlerentdeckungszeiten bezogen auf die farbliche Gestaltung.....	115
Abbildung 5-28: Mittelwerte der sechs Dimensionen bezogen auf das Geschlecht	119
Abbildung 5-29: Mittelwerte der UX-Qualitäten auf das Geschlecht der Versuchsteilnehmer	120
Abbildung 5-30: Mittelwerte der sieben Items bezogen auf das Geschlecht	121
Abbildung 5-31: Mittelwerte der Differenzen der Dimensionen bezogen auf das Geschlecht.....	124
Abbildung 5-32: Mittelwerte <i>Gute-Schlechte-Stimmung</i> bez. auf das Geschlecht.....	125
Abbildung 5-33: Mittelw. positiver Aspekt <i>Gute-Schlechte-Stimmung</i> bez. auf das Geschlecht.....	125
Abbildung 5-34: Mittelw. negativer Aspekt <i>Gute-Schlechte-Stimmung</i> bez. auf das Geschlecht	126

Abbildung 5-35: Mittelwerte <i>Wachheit-Müdigkeit</i> bez. auf das Geschlecht.....	126
Abbildung 5-36: Mittelw. positiver Aspekt <i>Wachheit-Müdigkeit</i> bez. auf das Geschlecht.....	127
Abbildung 5-37: Mittelw. negativer Aspekt <i>Wachheit-Müdigkeit</i> bez. auf das Geschlecht	127
Abbildung 5-38: Mittelwerte <i>Ruhe-Unruhe</i> bez. auf das Geschlecht.....	128
Abbildung 5-39: Mittelw. positiver Aspekt <i>Ruhe-Unruhe</i> bez. auf das Geschlecht	129
Abbildung 5-40: Mittelw. negativer Aspekt <i>Ruhe-Unruhe</i> bez. auf das Geschlecht	129
Abbildung 5-41: Mittelwerte Messpunkte für Bedienungssicherheit bez. auf das Geschlecht.....	130
Abbildung 5-42: Mittelwerte Messpunktdifferenzen bezogen auf das Geschlecht	131
Abbildung 5-43: Mittelwerte Fehlerentdeckungszeiten bezogen auf das Geschlecht	132
Abbildung 5-44: Mittelwerte Dimension <i>Attraktivität</i> bezogen auf die Wechselwirkung.....	135
Abbildung 5-45: Mittelwerte Dimension <i>Durchschaubarkeit</i> bezogen auf die Wechselwirkung.....	136
Abbildung 5-46: Mittelwerte Dimension <i>Effizienz</i> bezogen auf die Wechselwirkung.....	136
Abbildung 5-47: Mittelwerte Dimension <i>Steuerbarkeit</i> bezogen auf die Wechselwirkung.....	137
Abbildung 5-48: Mittelwerte Dimension <i>Stimulation</i> bezogen auf die Wechselwirkung.....	137
Abbildung 5-49: Mittelwerte Dimension <i>Originalität</i> bezogen auf die Wechselwirkung.....	138
Abbildung 5-50: Mittelwerte des Items 4 bezogen auf die Wechselwirkung.....	138
Abbildung 5-51: Mittelwerte des Items 7 bezogen auf die Wechselwirkung.....	139
Abbildung 5-52: Mittelwerte des Items 5 bezogen auf die Wechselwirkung.....	139
Abbildung 5-53: Mittelwerte des Gesamtsicherheitszuwachses bez. auf die Wechselwirkung.....	141
Abbildung 5-54: Mittelwerte des Messpunkts 4 bezogen auf die Wechselwirkung.....	141
Abbildung 5-55: Mittelwerte Unregelmäßigkeit <i>11_SS_S23</i> bezogen auf die Wechselwirkung	142
Abbildung 5-56: Mittelwerte Unregelmäßigkeit <i>17_GFM_W02</i> bezogen auf die Wechselwirkung ..	143
Abbildung 5-57: Mittelwerte Unregelmäßigkeit <i>18_ZLB_S23</i> bezogen auf die Wechselwirkung	143
Abbildung 5-58: Mittelwerte Unregelmäßigkeit <i>19_ZLB_S22</i> bezogen auf die Wechselwirkung	144
Abbildung 5-59: Mittelwerte Unregelmäßigkeit <i>20_ZLB_S23</i> bezogen auf die Wechselwirkung	144
Abbildung 5-60: Mittelwerte Unregelmäßigkeit <i>25_ZLB_S01</i> bezogen auf die Wechselwirkung	145
Abbildung 5-61: Übersicht der Abhängigkeiten (ohne Darstellung der Wirkrichtung).....	158

TABELLENVERZEICHNIS

Tabelle 2-1: Die sieben Grundsätze der Dialoggestaltung	31
Tabelle 2-2: Stufen der Konformität der WCAG 2.0.....	40
Tabelle 2-3: Gegenüberstellung der Vor- und Nachteile der betrachteten Darstellungsarten	44
Tabelle 3-1: Übersicht der Leistungskennzahlen	51
Tabelle 4-1: RGB-Wert der einzelnen Bildelemente	59
Tabelle 4-2: Übersicht über die relative Farbhelligkeit und das Kontrastverhältnis.....	60
Tabelle 4-3: Übersicht über die Helligkeitsdifferenz und den Farbunterschied	61
Tabelle 4-4: Übersicht der Hilfshandlungen bei einer Signalstörung.....	65
Tabelle 4-5: Übersicht der Hilfshandlungen bei einer Gleisfreimeldestörung.....	66
Tabelle 5-1: Fragebogen zur Beurteilung des Trainings	76
Tabelle 5-2: Auswahl an Fragebögen zur Skalierung der Befindlichkeit	78
Tabelle 5-3: Interpretation der Dimensionen des MDBF	80
Tabelle 5-4: Fragebogen zur Beurteilung der Bedienungssicherheit	81
Tabelle 5-5: Übersicht über die gezählten Hilfshandlungen	89
Tabelle 5-6: Übersicht über die Klassifizierung der Bedienfehler	90
Tabelle 5-7: Darstellung des Versuchsablaufs.....	91
Tabelle 5-8: Gegenüberstellung der verwendeten statistischen Verfahren	95
Tabelle 5-9: Übersicht über die Antwortkategorie des Freitextitems 8.....	101
Tabelle 5-10: Übersicht der Antwortkategorien des Freitextitems 9	103
Tabelle 5-11: Mittelw. und Standardabw. der Fehlerentdeckungszeit (farbliche Gestaltung).....	117
Tabelle 5-12: Interferenzstatistische Ergebnisse für die vier Messpunkte	130
Tabelle 5-13: Mittelw. und Standardabw. der Fehlerentdeckungszeit (Geschlecht).....	133
Tabelle 5-14: Übersicht der aufgestellten Hypothesen	146
Tabelle A-1: Auswertung User Experience (Test auf Normalverteilung)	194
Tabelle A-2: Auswertung User Experience (t-Test für unabhängige Stichproben)	194
Tabelle A-3: Auswertung User Experience (U-Test)	195
Tabelle A-4: Auswertung User Experience (Störungsbewertungen).....	195
Tabelle A-5: Auswertung Beanspruchung (Test auf Normalverteilung)	196
Tabelle A-6: Auswertung Beanspruchung (t-Test für unabhängige Stichproben)	196
Tabelle A-7: Auswertung Beanspruchung (U-Test)	197

Tabelle A-8: Auswertung Beanspruchung (Störungsbewertungen).....	197
Tabelle A-9: Auswertung Befindlichkeit (Test auf Normalverteilung)	198
Tabelle A-10: Auswertung Befindlichkeit (t-Test für unabhängige Stichproben)	199
Tabelle A-11: Auswertung Befindlichkeit (t-Test für abhängige Stichproben).....	199
Tabelle A-12: Auswertung Befindlichkeit (t-Test für abhängige Stichproben).....	199
Tabelle A-13: Auswertung Befindlichkeit (U-Test)	200
Tabelle A-14: Auswertung Befindlichkeit (Störungsbewertungen).....	200
Tabelle A-15: Auswertung Bedienungssicherheit (Test auf Normalverteilung).....	201
Tabelle A-16: Auswertung Bedienungssicherheit (t-Test für unabhängige Stichproben).....	201
Tabelle A-17: Auswertung Bedienungssicherheit (U-Test)	202
Tabelle A-18: Auswertung Bedienungssicherheit (Störungsbewertungen)	202
Tabelle A-19: Auswertung Leitungskennzahlen I (Test auf Normalverteilung).....	203
Tabelle A-20: Auswertung Leitungskennzahlen II (Test auf Normalverteilung).....	203
Tabelle A-21: Auswertung Leitungskennzahlen III (Test auf Normalverteilung).....	203
Tabelle A-22: Auswertung Leistungskennzahlen (t-Test für unabhängige Stichproben)	204
Tabelle A-23: Auswertung Leistungskennzahlen (U-Test).....	204
Tabelle A-24: Auswertung Leitungskennzahlen (Störungsbewertungen)	205

1 EINLEITUNG

1.1 Motivation

Die Fehleranfälligkeit des Menschen am Arbeitsplatz zu verringern, ist eine zentrale Zielsetzung aller Maßnahmen zur Erhöhung der Sicherheit¹. Um dies zu erreichen, wurden in der Vergangenheit vermehrt technische Systeme eingesetzt. Dies führte zu einer zunehmenden Automatisierung von Arbeitsprozessen, was neben den unbestreitbaren wirtschaftlichen Vorteilen und einer Erhöhung der Sicherheit, auch mit Risiken behaftet ist [Fab09]. Auch im Eisenbahnwesen ist diese Entwicklung zu erkennen. Am Fahrdienstleiterarbeitsplatz beispielsweise, ermöglicht die zunehmende Automatisierung eine Vergrößerung der Steuerbereiche. Hierdurch ergeben sich neuartige Herausforderungen für die Entwicklung von Bedienplätzen, welche Auslöser für die Überlegungen und Untersuchungen der vorliegenden Arbeit sind.²

Durch die größer werdenden Steuerbereiche wird dem Fahrdienstleiter (Fdl) die gesamthafte Erfassung der Betriebssituation erschwert. Gerade in Betriebszentralen führt das zu Problemen, da dadurch die fahrdienstlichen Fehlhandlungen zunehmen können [Pac16]. Zudem führt der hohe Automatisierungsgrad zu einer Verlagerung des Aufgabenspektrums am Fahrdienstleiterarbeitsplatz. Regelbasierte Routinetätigkeiten verlieren zusehends an Bedeutung, weil sie durch technische Systeme übernommen werden. Das Ergebnis dieser Entwicklung ist, dass der Fdl während des störungsfreien Regelbetriebs aus der aktiven Betriebsführung ausgeschlossen wird und eine passive Rolle einnimmt. Im Regelbetrieb stellt das kein Problem dar, im Störfall kann dies jedoch zu einer nicht vernachlässigbaren Herausforderung werden. Hier wird der Fdl wieder zu einem integralen Bestandteil der Handlungskette, wobei er sich gleichzeitig mit einer sehr hohen Arbeitsbelastung konfrontiert sieht.

Wie nicht erst das Ereignis in Bad Aibling zeigt [BR16], hat der Mensch einen nicht unerheblichen Einfluss auf die Sicherheit im Eisenbahnwesen und Unachtsamkeit kann schnell verheerende Folgen haben. Da es bis zum vollautomatisierten Zugbetrieb noch ein langer Weg ist, bleibt der Mensch vorerst ein essentieller Bestandteil im Eisenbahnwesen. Aus diesem Grund ist es unerlässlich seine Arbeitsmittel der neuen Situation entsprechend anzupassen.

¹ „Freiheit von unvermeidbaren Risiken für die Gesundheit des Menschen oder für die Umwelt.“ ([DIN16], S. 20)

² Die Beschränkung auf den Fahrdienstleiterarbeitsplatz wurde bewusst gewählt. Das Aufgabenspektrum des Fahrdienstleiters hat sich in den letzten Jahrzehnten sehr stark gewandelt. Des Weiteren hat der Fahrdienstleiter mit seiner Arbeit einen direkten Einfluss auf die Abwicklung des Eisenbahnbetriebs.

Das ist auch ein Grund, warum es seit geraumer Zeit die Bestrebung gibt, einen neuen, einheitlichen Bedienplatz zu entwickeln [BaBl14]. Bei der Neuentwicklung dieses Bedienplatzes stehen jedoch fast nur funktionale Aspekte im Vordergrund. Dass diese Betrachtung zu kurz greift und menschliche Fähigkeiten sowie persönliche Bedürfnisse ebenso betrachtet werden müssen, soll mit der vorliegenden Arbeit gezeigt werden. Gerade das Einbeziehen der eben genannten „weichen“ Faktoren erfolgte in der Vergangenheit nur wenig oder wenig zielgerichtet – gerade im Eisenbahnwesen.

Um diese Lücke zu schließen, wird im Rahmen dieser Arbeit eine Studie an der Stellwerksbedienoberfläche³ durchgeführt. Das Hauptaugenmerk der Studie liegt auf der Untersuchung des Einflusses von verschiedenen Bedienoberflächendesigns auf den Bediener. Vor allem die beiden Aspekte *Farbe* und *Kontrast*⁴ spielen hierbei eine entscheidende Rolle. Der Ausgangspunkt der Überlegungen ist zum Teil auf Erkenntnisse aus anderen Industriezweigen zurückzuführen. So wurde beispielsweise festgestellt, dass bestimmte Farben eingesetzt werden können, um Aufmerksamkeit zu erregen [OsDa85] und die empfundene Ästhetik des interaktiven Systems einen Einfluss auf die Bewertung der Usability hat [TKI00]. Der letztgenannte Effekt ist besonders von Interesse, da sich eine gute Usability in einer höheren Akzeptanz des Benutzers gegenüber dem System niederschlagen kann [Mos12]. Das hat zur Folge, dass dessen Leistungsbereitschaft erhöht werden kann [MoTh10] und Bedienfehler verringert werden können [VaDe02].

Ob diese Erkenntnisse auch auf das Eisenbahnwesen übertragen werden können und ob „weiche“ Faktoren bei der Entwicklung neuer Bedienoberflächen betrachtet werden müssen, wird in der vorliegenden Bedienoberflächenstudie überprüft.

1.2 Zielsetzung

Da bei der Entwicklung neuer Bedienoberflächen oft nur funktionale Aspekte betrachtet werden, ist das Ziel der durchgeführten Studie, den Einfluss der farblichen Gestaltung der Bedienoberfläche eines Stellwerks auf den Bediener zu untersuchen. Neben funktionalen Aspekten (z. B. der Usability), werden in der Studie auch „weiche“ Faktoren (z. B. die Befindlichkeit) betrachtet.⁵ Die gewonnenen Erkenntnisse sollen verwendet werden können, um Vorgaben für zukünftige Bedienoberflächen abzuleiten.

Um den Einfluss der farblichen Gestaltung näher beschreiben zu können, werden zwei Bedienoberflächen einer deutschen Stadtbahnsimulation miteinander verglichen. Diese unterscheiden sich nur hinsichtlich der farblichen Gestaltung, nicht aber in Bezug auf Benutzbar-

³ Die Bedienoberfläche ist ein Bestandteil der Benutzerschnittstelle im Stellwerk und bildet, neben den Bedienelementen, das Bindeglied zwischen FdI und Stellwerk.

⁴ im weiteren Verlauf als *farbliche Gestaltung* bezeichnet

⁵ Dieses Vorgehen wird beispielsweise durch das *Canadian Centre for Occupational Health and Safety* (CCOHS) unterstützt [CCOHS14].

keit und Funktionalität voneinander. Im Rahmen des Experiments wird die Wirkung der Bedienoberfläche auf die folgenden Variablen untersucht:

- User Experience,
- psychische Beanspruchung,
- Befindlichkeit,
- Bedienungssicherheit und
- diverse Leistungskennzahlen.

Mithilfe der erfassten Leistungskennzahlen soll ein theoretischer Ansatz entwickelt werden, um die Bedienereffizienz am Fahrdienstleiterarbeitsplatz versuchsweise zu operationalisieren. Dabei liegt der Fokus der Arbeit ausschließlich auf der Konstruktion des Konzepts und der statistischen Überprüfung der theoretisch beschriebenen Zusammenhänge.

Ein weiterer Bestandteil der vorliegenden Arbeit ist die Suche nach einem Konzept, mit dessen Hilfe die Bedienoberfläche eines Stellwerks auf ihre Benutzerfreundlichkeit überprüft werden kann. Hintergrund dieser Zielstellung ist die Forderung nach einem einfachen Verfahren, mit dessen Hilfe eine erste Näherung möglich ist, ohne aufwendige Analysen durchführen zu müssen. Die Ergebnisse des Konzepts werden dann im Rahmen der Auswertung mit den gewonnenen Daten der Studie verglichen, um eine Aussage über die Gültigkeit der Näherung treffen zu können.

1.3 Vorgehensweise

Zu Beginn der Arbeit werden einige Grundlagen für das bessere Verständnis der Aufgabenstellung aufbereitet. Hierzu werden die veränderten Arbeitsbedingungen des Fdl und die damit verbundenen Probleme aufgezeigt (Kapitel 2.1.1). Im weiteren Verlauf wird auf mögliche Ursachen für menschliche Fehler eingegangen und ein Exkurs über die Klassifizierung eben dieser gegeben (Kapitel 2.1.2). Diese Beschreibung ist notwendig, um die Ursprünge besser zu verstehen. Zudem wird in Kapitel 2.1.3 der Begriff *Human Factors* eingeführt, erläutert und auf das Eisenbahnwesen übertragen.

Die Lösung zur Reduzierung menschlicher Fehler scheint eine frühzeitige Einbindung des Menschen in die Systementwicklung zu sein. Um diese Forderung besser nachvollziehen zu können, werden in Kapitel 2.2.1 die Benutzerfreundlichkeit definiert, einige Grundlagen bzgl. des Zusammenspiels zwischen Mensch und Maschine vermittelt und die Besonderheiten der Mensch-Maschine-Interaktion aufgezeigt. In diesem Zusammenhang werden zwei Wissenschaften vorgestellt, die sich mit der Erforschung der Mensch-Maschine-Interaktion befassen und dadurch zu einem besseren Verständnis in Bezug auf die Benutzerfreundlichkeit beitragen. Danach werden zwei Ansätze näher erläutert, welche sich für die Bewertung der Benutzerfreundlichkeit eignen (Kapitel 2.2.2). Im weiteren Verlauf werden zwei Attribute

vorgestellt, welche einen wesentlichen Einfluss auf die Bewertung der Benutzerfreundlichkeit haben (Kapitel 2.3.1). In Kapitel 2.3.2 werden Mindestanforderungen und Bewertungsmaßstäbe gesucht, um diese Attribute bei der Entwicklung neuer Bedienoberflächen einbeziehen zu können. Danach werden in Kapitel 2.4.1 die grundsätzlichen Darstellungsarten vorgestellt. Da sich diese Unterscheidung für die verwendeten Bedienoberflächen nicht eignet, wird ein neuer Ansatz eingeführt (Kapitel 2.4.2). Um die Vorbetrachtungen abzuschließen, wird in Kapitel 2.4.3 die aktuelle Situation bei der DB Netz AG kurz beleuchtet.

Anschließend wird die Zielstellung der Arbeit beschrieben (Kapitel 3) und die Versuchsumgebung vorgestellt (Kapitel 4). Die Beschreibung der Versuchsumgebung beinhaltet die Vorstellung der verwendeten Bedienplätze (Kapitel 4.2.1/Kapitel 4.2.2) und die Erklärung der Ausleuchtung (Kapitel 4.2.3). In Kapitel 4.2.4 wird der neue Ansatz zur Einteilung der Darstellungsart der Bedienoberfläche aus Kapitel 2.4.2 auf die verwendeten Bedienplätze übertragen. Zum Schluss werden die Unregelmäßigkeiten, die während des Versuchs eintreten, vorgestellt und die notwendigen Bedienhandlungen erläutert (Kapitel 4.3).

Kapitel 5 bildet den Hauptteil der Arbeit und beschreibt die Bedienoberflächenstudie im Detail. Das Versuchskapitel gliedert sich in drei Teile. Im ersten Teil wird die Methode beschrieben (Kapitel 5.1). Hierzu zählen die Stichprobe (Kapitel 5.1.1), der Versuchsplan (Kapitel 5.1.2) und das verwendete Material (Kapitel 5.1.3). In Kapitel 5.1.4 wird der theoretische Ansatz näher beschrieben, mit dem die Bedienereffizienz im Stellwerk versuchsweise operationalisiert werden soll und in Kapitel 5.1.5 die Durchführung der Bedienoberflächenstudie. Den Abschluss bildet Kapitel 5.1.6. Das Kapitel schafft die statischen Grundlagen und gibt einen Überblick über die verwendeten Testverfahren. In Kapitel 5.2 erfolgt die Darstellung der Ergebnisse bezogen auf die farbliche Gestaltung (Kapitel 5.2.1), das Geschlecht (Kapitel 5.2.2) und die Wechselwirkung zwischen beiden unabhängigen Variablen (Kapitel 5.2.3). Die auf die Hypothesen bezogene Diskussion der Ergebnisse, findet in Kapitel 5.3 statt.

Am Ende der Arbeit (Kapitel 6) werden die Erkenntnisse übersichtlich zusammengefasst und es wird ein Ausblick auf mögliche weiterführende Untersuchungen gegeben.

2 VORBETRACHTUNGEN UND HINTERGRÜNDE

2.1 Mensch und Technik

Sicherheit und Zuverlässigkeit sind zwei Attribute, die für das Eisenbahnwesen von zentraler Bedeutung sind. Seit die erste Eisenbahn im September 1825 in England ihren Betrieb aufnahm [Pot88], werden Betriebsverfahren und Sicherungseinrichtungen immer wieder auf Neue ergänzt und verbessert. Oft geschieht dies mit dem Ziel, den Menschen zu ersetzen. Die Folge ist ein wachsender Automatisierungsgrad, der neben den unbestreitbaren wirtschaftlichen Vorteilen und einer Erhöhung der Sicherheit und Zuverlässigkeit, auch mit Risiken behaftet ist [Fab09].

Ziel dieses Kapitels ist es, einige der wichtigsten Themen im Zusammenspiel von Mensch und Maschine vorzustellen und die Rolle der menschlichen Einflussfaktoren innerhalb des Eisenbahnsystems zu veranschaulichen. Ferner schafft das Kapitel den Rahmen, der für das Verständnis der weiteren Überlegungen notwendig ist.

2.1.1 Einführung

Der Fahrdienstleiter ist für die sichere und planmäßige Durchführung des gesamten Bahnverkehrs innerhalb seines Zuständigkeitsbereiches verantwortlich. Er besitzt somit eine zentrale Rolle im Eisenbahnwesen. Aufgrund zunehmender Technisierung haben sich Bedienelemente, Anzeigen und das Aufgabenspektrum des Fdl vollständig gewandelt. Ist im mechanischen Stellwerk der Großteil der Handlungen durch sensomotorische Fähigkeiten und menschliche Muskelkraft geprägt, rücken bei der Bedienung eines modernen Stellwerks überwachende Tätigkeiten und kognitive Prozesse in den Vordergrund. Nachfolgend wird ein Überblick über die Anfänge und die Entwicklung der Sicherungstechnik gegeben und es wird auf mögliche Schwächen automatisierter Systeme eingegangen.

2.1.1.1 Die Entwicklung des Fahrdienstleiterarbeitsplatzes

Der Schienenverkehr hat den systemimmanenten Vorteil, dass er sich besonders gut zum Transport großer Massen mit hohen Geschwindigkeiten eignet. Grund dafür sind zwei wesentliche Systemeigenschaften, die dies ermöglichen und den Schienenverkehr vom Stra-

ßenverkehr unterscheiden. Die Rede ist von der schienengebundenen Spurführung und der geringen Haftreibung zwischen Fahrzeug und Infrastruktur. Diese Stärken des Schienenverkehrs bergen gleichzeitig Gefahren. So macht die Spurführung den Fahrtrichtungswechsel aus dem Triebfahrzeug heraus unmöglich. Ausweichen und Überholen ist ohne die Bereitstellung einer entsprechenden Infrastruktur undenkbar. Hierfür wurden Weichen entwickelt. Das zweite Problem entsteht durch die bewegten Massen. Große Masse, hohe Geschwindigkeit und geringe Haftreibung führen zu sehr langen Bremswegen. Diese Bremswege können von dem Triebfahrzeugführer (Tf) nicht eingesehen werden und müssen daher durch besondere Techniken zur Steuerung und Sicherung des Fahrwegs sowie zur Regelung der Zugfolge bereitgestellt und überwacht werden [Pac13]. Dazu gehören Signale.

Weichen und Signale wurden in der Frühzeit durch Handstelleinrichtungen vor Ort bedient. Durch den wachsenden Verkehr stellte sich im Laufe der Zeit die Notwendigkeit einer mechanischen Abhängigkeit zwischen den einzelnen Elementen heraus. Abhilfe wurde durch den englischen Eisenbahner *JOHN SAXBY* mit seinem Patent im Jahre 1856 geschaffen. Mit der Entwicklung des Zentralapparates (später als Stellwerk bezeichnet) wurden die einzelnen Bedienelemente für Weichen und Signale zentralisiert und ab 1859 durch die Entwicklung eines Verschlusssystems in Abhängigkeit gebracht [Pot88]. Durch die Sicherheitslogik zwischen den zu steuernden Elementen konnten Gefährdungen minimiert werden. Infolgedessen war der Siegeszug der Signaltechnik nicht mehr aufzuhalten.⁶

Obgleich sich im Anlagenbestand der DB Netz AG noch ein sehr großer Anteil (73 %) von mechanischen Stellwerken und von Drucktastenstellwerken befindet [DBAG13a], nimmt der Anteil an elektronischen Stellwerken (ESTW) zu. Das Problem an dem aktuellen Modernisierungskonzept der DB Netz AG ist, dass die Erneuerung der Alttechniken, durch den Bau eines ESTW, mit sehr langen Planungszeiten und hohen Investitionskosten verbunden ist. Hinzu kommt der kurze Produktlebenszyklus der proprietären elektronischen Komponenten. Das hat zur Folge, dass die alte Technik nicht schnell genug erneuert werden kann und ein Modernisierungsstau entsteht. Um dem entgegenzuwirken, verfolgt die DB Netz AG seit einigen Jahren die Standardisierung der technischen Schnittstellen im Projekt *Neuausrichtung Produktionssteuerung* (NeuPro). Integraler Bestandteil von NeuPro ist die Entwicklung des digitalen Stellwerks (DSTW). Als bundesweit erster Pilot kommt die neue Stellwerkstechnik im Regionalnetz Harz-Weser zum Einsatz. Vorteile des DSTW sind neben den standardisierten, offenen Schnittstellen, eine moderne Netzwerkarchitektur zur Datenübertragung und eine dezentrale Energieversorgung [DBAG15]. Durch die Trennung von Energie und Daten entfällt beispielsweise die Limitierung der Stellentfernung. Im Zuge dieses Pilotprojekts soll auch ein neues Bedienplatzsystem erprobt werden – das integrierte Bedienplatzsystem (iBS). Mit dem iBS wird eine systemunabhängige und hardwareneutrale Bedienumgebung für LST, Disposition, Meldeanlagen und Telekommunikation geschaffen [DBAG15].

⁶ Eine Übersicht über die Anfänge der Signaltechnik gibt *SASSE* mit seinem Beitrag *Streifzug durch die Geschichte der deutschen Signaltechnik* in einem Signal & Draht Sonderdruck aus den 50er Jahren [Sas58].

Ein weiteres Argument für NeuPro ist, dass die gewonnenen Erkenntnisse und Teile der neuen Technik im europäischen Projekt *European Initiative Linking Interlocking Subsystems* (EULYNX) weiterverwendet werden können. EULYNX ist eine Idee europäischer Infrastrukturbetreiber zur Standardisierung der Stellwerkstechnik auf internationaler Ebene [NN16b].

Die Einführung der ESTW in der Vergangenheit und der DSTW in der Zukunft führten bzw. führen dazu, dass sich die Bedienelemente, die Anzeigen und das Aufgabenspektrum des Fdl vollständig gewandelt haben und noch weiter wandeln werden. Um die zunehmende Komplexität der Anlagen in den Griff zu bekommen, nimmt auch der Automatisierungsgrad zu. Die Automatisierung ermöglicht in diesem Zusammenhang eine zentrale Steuerung und Überwachung größerer Bereiche, wodurch die Sicherheit erhöht werden soll. Das stimmt nur bedingt, denn durch die Zentralisierung wird qualifiziertes Personal reduziert. Das hat u. a. zur Folge, dass die Ortskenntnis je Fahrdienstleiter abnimmt und sich die Anzahl der registrierungspflichtigen Hilfshandlungen⁷ durch die größeren Steuerbereiche erhöht. Dass das nicht die einzigen Schwächen der Automatisierung sind, wird nachfolgend gezeigt.

2.1.1.2 Schwächen der Automatisierung

Im mechanischen und elektromechanischen Stellwerk erfolgt die Visualisierung der Betriebsituation über die Hebelstellung der Feldelemente und die Anzeige der wirksamen Verschlüsse [Pac98]. Dadurch hat der Fdl die Möglichkeiten den Status seiner Anlage direkt abzufragen und detaillierte Rückmeldungen zu erhalten. Beispielsweise kann der Fdl sich im mechanischen Stellwerk Veränderungen der Außenanlage vor Ort anschauen oder über den erhöhten Kraftaufwand beim (Um)Stellen des Feldelementes frühzeitig erkennen. Dieses Beispiel zeigt, dass der Fdl aktiv in den Sicherungsprozess eingebunden ist. Das heutige Betriebskonzept sieht die Zentralisierung in sieben Betriebszentralen vor, wodurch sich die Steuerbereiche ausdehnen. Um mit der wachsenden Informationsdichte am Arbeitsplatz umgehen zu können, werden immer mehr Aufgaben durch technische Lösungen übernommen. Ein modernes Stellwerk (ESTW und DSTW) hält den Betrieb selbsttätig aufrecht und alarmiert den Bediener⁸ im Allgemeinen nicht, bevor eine Unregelmäßigkeit auftritt. Die geänderte Situation führt zu der Frage, welche Probleme durch die Verlagerung des Tätigkeitsspektrums entstehen. Um diese Frage beantworten zu können, wird zunächst erklärt, was unter einem *automatisierten System* zu verstehen ist. Im Anschluss daran werden die Schwächen solcher Systeme aufgezeigt.

Ein automatisiertes System stellt das Resultat eines Automatisierungsprozesses dar. Dieser Prozess besteht daraus, einzelne Funktionen oder ganze Tätigkeiten auf eine Maschine zu verlagern [HaTi00]. Das Tätigkeitsspektrum des Fahrdienstleiters hat sich mit der Zeit von

⁷ Hilfs- und Ersatzhandlungen, bei denen der Bediener unter Umgehung der Sicherungsebene des Stellwerks unmittelbar Sicherheitsverantwortung übernimmt. Diese müssen für nachträgliche Auswertungen registriert werden [NaPa04]. Die Registrierung erfolgt in der Regel über ein Zählwerk.

⁸ Der Begriff des Bedieners umfasst alle Personen, die mit der Bedienung von Signalanlagen vertraut sind [DBAG13b]. Der Bediener kann, muss aber kein Fahrdienstleiter sein.

einer Eigentätigkeit zu einer reinen Überwachungstätigkeit⁹ verschoben. Das hat zur Folge, dass der Fdl den Regelkreis verlässt und vom eigentlichen Sicherungsprozess ausgeschlossen ist, da dieser weitgehend verdeckt abläuft – *supervisory control* [She87]. Mit dem Ausschluss aus dem Regelkreis sollen seine Schwächen kompensiert und die Systemsicherheit und Zuverlässigkeit erhöht werden [She87]. Dass dies nicht immer zutreffend ist, wird bereits in den 70er Jahren des letzten Jahrhunderts am Beispiel der Cockpitautomatisierung in der Luftfahrt gezeigt. Die zunehmende Unterstützung durch die Technik führt zu einem Fertigkeitsverlust, welcher im Störfall zum Problem wird [ShJo76]. Studien im Flugsimulator zeigen, dass Piloten bei einem gezwungenen Übergang vom Autopilot auf manuelle Bedienung eher zu Fehlern neigen, als wenn sie von Anfang an selbst geflogen wären [JPS76].

Abbildung 2-1 zeigt die vereinfachte Systemstruktur eines ESTW – bestehend aus Zuglenkung und Sicherungstechnik – und die Rolle des Fdl. Im Regelbetrieb wird der Fdl von Routineaufgaben entlastet, indem das ESTW die Information selbstständig verarbeitet, Abweichungen vom Soll regelt und das gewünschte Ergebnis erzeugt. Fällt eine technische Komponente aus oder kann aus betrieblichen Gründen nicht mehr genutzt werden, wird der Betrieb in der Rückfallebene¹⁰ durchgeführt. Der Fdl muss die automatisierten Funktionen wieder manuell übernehmen und/oder mithilfe von betrieblichen Maßnahmen den Betrieb aufrechterhalten. Dieser Eingriff wird über den Schalter in Abbildung 2-1 symbolisiert.

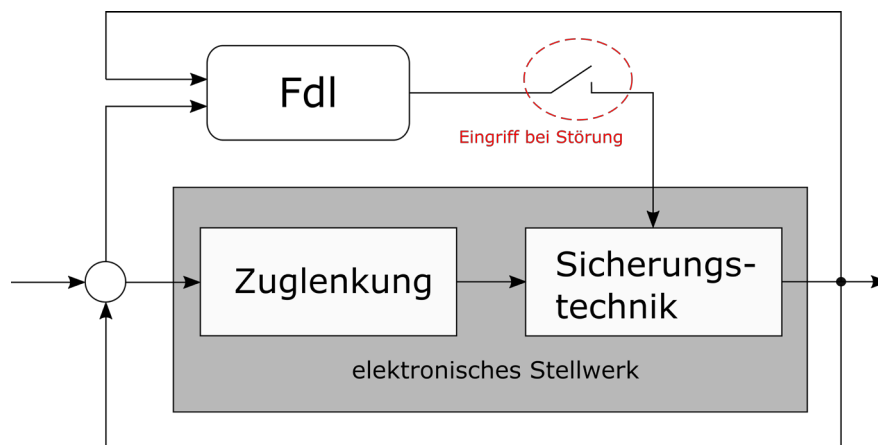


Abbildung 2-1: Darstellung der vereinfachten Systemstruktur eines ESTW mit der Rolle des Fdl (in Anlehnung an BUBBS monitives System [Bub92])

Hatten Fehlhandlungen des Fdl in der Vergangenheit ihre Ursache zumeist in unzureichender Kenntnis oder nachlässiger Anwendung des Regelwerks, sind im ESTW vermehrt Fehleinschätzungen der Betriebssituation Ursache für fahrdienstliche Fehlhandlungen [Pac16]. Un-

⁹ Eine Überwachungstätigkeit liegt vor, wenn der Mensch die Aufgabe hat, das Funktionieren einer Anlage fortlaufend zu überprüfen und ggf. korrigierend einzugreifen [Sch81b].

¹⁰ Da auch in hochverfügbaren Systemen mit Störungen zu rechnen ist, muss eine Rückfallebene vorgesehen werden. Hierbei handelt es sich um einen definierten Betriebszustand, der nicht dem Regelbetrieb entspricht. Der Regelbetrieb impliziert die volle Funktion aller technischen Systeme im Bahnbetrieb. Die Rückfallebene unterscheidet sich vom Regelbetrieb durch die eingeschränkte Leistungsfähigkeit bei gleichem Verkehrsaufkommen und die erhöhte Beanspruchung des Betriebspersonals [Wie07]. Durch eine hohe Systemverfügbarkeit wird diese unerwünschte Form des Betriebs minimiert.

zureichende Rückmeldungen und fehlende Transparenz des Systems sowie ein zu hohes Vertrauen in die Anlage begünstigen diesen Umstand. *WICKENS UND HOLLANDS* fassen die negativen Folgen dieser Entwicklung unter dem Begriff *out-of-the-loop-unfamiliarity* (OOTLUF) zusammen [WiHo00]. Die Autoren sehen drei Problemfelder der OOTLUF, welche im Folgenden näher beschrieben werden.

2.1.1.2.1 Abnahme des Situationsbewusstseins

Die Fülle an Untersuchungen zur Auswirkung überwachender Aufgaben auf das Arbeitsverhalten des Menschen werden mit einer einfachen Feststellung zusammengefasst: „*However, as the long history of vigilance research can document, humans do not make terribly good monitors.*“ ([Wic99], S. 225). Kurz, der Mensch eignet sich nur bedingt als Überwacher eines technischen Systems. Es fällt ihm schwer, seine Konzentration und Aufmerksamkeit unter monotonen Bedingungen aufrechtzuerhalten. Dieser Effekt verstärkt sich mit zunehmendem Alter und unzureichendem Training [Kor14]. Das widerspricht jedoch der Entwicklung der letzten Jahre, denn die Hauptaufgabe des Fdl im modernen Stellwerk ist die Überwachung einer in sich geschlossenen sicherungstechnischen Anlage (Abbildung 2-1). Anstatt ein aktiver Teil des Regelkreises zu sein, übernimmt er eine eher passive Rolle. Damit verbunden ist eine verminderte aktive Auseinandersetzung mit dem System. Diese hat das Verlernen von Fertigkeiten (Kapitel 2.1.1.2.2) und den Verlust eines angemessenen Situationsbewusstseins zur Folge.

Zur Beschreibung des Situationsbewusstseins (*Situation Awareness, SA*) eignet sich das Modell von *ENDSLEY*. Die Autorin integriert die Wahrnehmung, die Aufmerksamkeit und das Gedächtnis in ein Modell der Informationsverarbeitung. *ENDSLEY* definiert Situationsbewusstsein wie folgt ([End88], S. 97):

“Situation awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future.”

Hieraus ergibt sich ein dreistufiger, hierarchisch aufgebauter Prozess (Abbildung 2-2). Grundvoraussetzung für eine angemessene Reaktion ist die Wahrnehmung der relevanten Informationen. Dies wird auf der ersten Ebene beschrieben. Ebene 2 beschreibt das Verstehen und das Verarbeiten der Informationen. Dem Bediener muss klar sein, welche Anforderungen aktuell an ihn gestellt werden. Im Anschluss daran muss der Bediener vorhersagen wie sich die Situation entwickelt, um seine Handlungen entsprechend anpassen zu können (Ebene 3). Zusammengefasst bedeutet das, dass ein angemessenes Situationsbewusstsein immer dann vorliegt, wenn der Mensch sich bei der Aufgabenerfüllung zu jedem Zeitpunkt darüber bewusst ist, was gerade passiert, warum es passiert und wie sich die Situation weiterentwickeln wird [Man08].

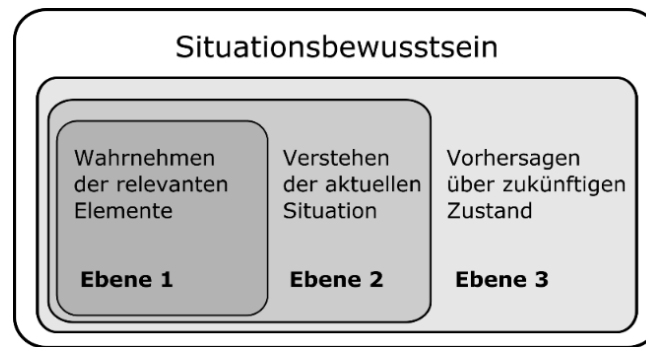


Abbildung 2-2: Die drei Ebenen des Situationsbewusstseins nach *ENDSLEY* [End95]

Das Problem in modernen Stellwerken ist, dass sich mangelndes Situationsbewusstsein vor allem dann zeigt, wenn die Sicherungstechnik ausfällt oder widererwartend nicht einwandfrei funktioniert. Im Regelbetrieb lässt sich nicht immer direkt feststellen, ob der FdI zu jedem Zeitpunkt die volle Kontrolle über das Stellwerk hat und ob er die Betriebssituation richtig einschätzt. Im schlimmsten Fall führt mangelndes Situationsbewusstsein zu einer Fehleinschätzung der Betriebssituation.

2.1.1.2.2 Fertigkeitsverlust

Die Nachteile durch den Ausschluss des Bedieners vom aktiven Regelungsprozess spiegeln sich nicht nur im fehlenden Situationsbewusstsein wider, sondern gehen oft mit einem Verlust manueller Fertigkeiten und einer mangelnden Vorstellung vom System einher [KMZ+14]. Relevant in diesem Zusammenhang ist, dass ein Erhalt motorischer Fertigkeiten auch ohne Training sehr viel besser gelingt als etwa der Erhalt kognitiver Fähigkeiten [Pat92]. Somit stellt der Fertigkeitsverlust insbesondere dann ein Risiko dar, wenn kognitive Funktionen vom Menschen auf die Maschine übertragen werden [Man08] – wie es bei automatisierten Systemen der Fall ist.

Die Überwachungstätigkeit und die verminderte aktive Auseinandersetzung mit dem System führen bei langanhaltender Beobachtungstätigkeit, welche eine hohe Aufmerksamkeit erfordert, zu einer wachsenden Unterforderung¹¹, einem Leistungsabfall und einem Übungsverlust. Dieser Umstand kann sich dann in einer höheren Fehleranfälligkeit des Bedieners in der Rückfallebene widerspiegeln.¹² In der Rückfallebene wird vom Bediener vorausgesetzt, dass er zu jedem Zeitpunkt in der Lage ist, die eigentlich automatisierten Funktionen wieder selbst zu übernehmen. Das zeigt die Ironie dahinter, denn der Mensch wird als schwächstes Glied der Kette gesehen und daher durch die Automatisierung ersetzt. Nichtsdestotrotz wird von ihm verlangt, die Anlage zu überwachen, im Störfall einzugreifen und ggf. die manuelle Bedienung zu übernehmen. *BAINBRIDGE* beschreibt die Ironien der Automatisierung ausführlich [Bai83].

¹¹ *HARGUTT UND KRÜGER* untersuchen die Gefahr der Unterforderung während Autobahnfahrten [HaKr01].

¹² Studien in der Gehirnforschung bestätigen diesen Zusammenhang. Das Gehirn gehört zu den Organen, welche bei der Bewältigung von Routineaufgaben schneller altern, in einen passiven Zustand verfallen und dadurch Leistungseinbußen verzeichnen [Kor14].

Ein Beispiel für die gerade beschriebene Ironie ist die automatische Zuglenkung im ESTW (Abbildung 2-1). Die Zuglenkung wirkt auf die Sicherungstechnik und übernimmt die Abwicklung der Zugfahrten. Dadurch wird die Belastung des Bedieners gesenkt und die Leistungsfähigkeit des Systems erhöht [BWS+12]. Das Problem daran ist, dass sich der Bediener kaum noch Gedanken über die Steuerung machen muss, weil systembedingte Reize ausbleiben. Dadurch sinkt seine Aufmerksamkeit und das Situationsbewusstsein nimmt in diesem Zusammenhang zwangsläufig ab. Dem Bediener wird durch die technische Unterstützung die Möglichkeit genommen, die Betriebssituation richtig zu erfassen.

2.1.1.2.3 Zu geringes oder übersteigertes Vertrauen in automatisierte Systeme

Ein weiterer negativer Effekt ist ein übermäßiges Vertrauen in das automatisierte System. Im Gegensatz zu einem zu geringen Vertrauen, was sich in der Regel in einer mangelnden Nutzung äußert, ist ein zu hohes Vertrauen mit erheblichen Risiken für die Sicherheit verbunden. In der Vergangenheit kam es beispielsweise durch „blindes Vertrauen“ des FdI in die Gleisfreimeldeanlage wiederholt zu Zusammenstößen. Ein Beispiel ist der Beinahe-Zusammenstoß zweier S-Bahnen im Mainzer Hauptbahnhof im August 2013 [Fra13]. Infolge eines nicht regelkonformen Sandens, wurde die S8 Richtung Frankfurt von der niederfrequenten Gleisfreimeldeanlage nicht mehr registriert und „verschwand“ auf der Weiche. Der FdI vertraute der Anlage ohne eine Statusmeldung der S-Bahn einzuholen. Der Unfall konnte nur verhindert werden, da ausreichend Sicht vorhanden war und die Einfahrtgeschwindigkeit der S-Bahn aus der Gegenrichtung gering war.

Die Probleme, die aus einem übersteigerten Vertrauen resultieren, werden unter den englischen Schlagwörtern *complacency* und *automation bias* zusammengefasst [MoIn00]. Bei *complacency* handelt es sich um ein Phänomen, was sich in einer unzureichenden Überwachung oder Überprüfung automatischer Systeme äußert [MaBa05]. Sie kann zum Übersehen kritischer Systemzustände führen [WHB+13]. Besitzt das System eine automatische Diagnoseeinrichtung, kann es zu zwei weiteren Fehlertypen kommen. Verlässt sich der Bediener zu sehr auf das System und übersieht eine kritische Situation, weil sie ihm durch die Diagnose nicht angezeigt wird, wird von *omission* Fehlern gesprochen. Wohingegen sogenannte *commission* Fehler entstehen, wenn der Bediener die Anweisungen des Systems blind befolgt und eine fehlerhafte Anweisung ausführt. Beide Fehlertypen beschreiben das Phänomen *automation bias*.

Der wesentliche Unterschied zwischen beiden Phänomenen besteht darin, dass *automation bias* im Gegensatz zu *complacency* einen Automatisierungsfehler voraussetzt (Abbildung 2-3). Solange das System fehlerfrei arbeitet, ergeben sich keine negativen Folgen aus *complacency* für den Bediener. Er wird in seinem Handeln bestärkt und verzichtet auf eine genauere Überprüfung und Überwachung. Kommt ein Automatisierungsfehler hinzu, kann es, bedingt durch die vorausgegangene Leichtfertigkeit des Bedieners, durch *omission* bzw. *commission* Fehlern zu einer Gefährdung kommen. Eine umfassendere Beschreibung beider

Phänomene findet sich in [Bah08]. Die Autorin überträgt in ihrer Dissertation *complacency* und *automation bias* auf den Anwendungsbereich der Prozesssteuerung.

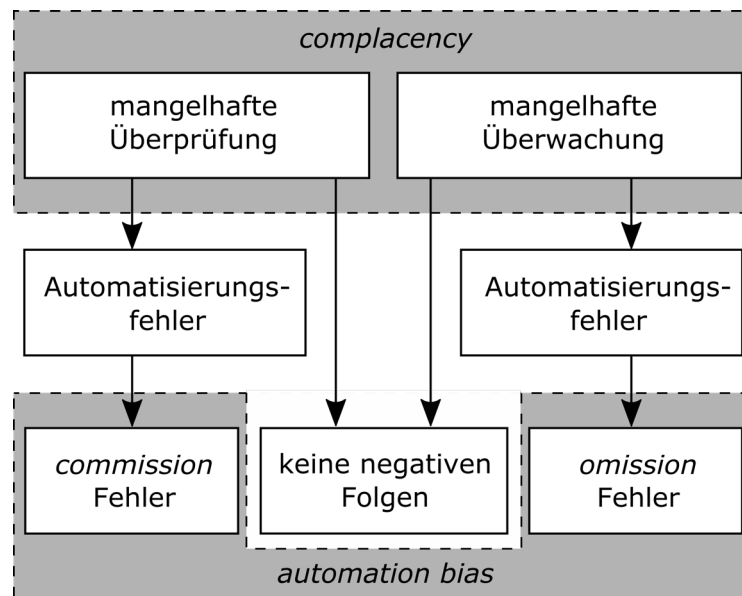


Abbildung 2-3: Zusammenhang der Konzepte *complacency* und *automation bias* (in Anlehnung an [Bah08])

2.1.2 Menschliche Fehler

Die vorangegangene Beschreibung zeigt, dass die Schwächen der Automatisierung teilweise durch grundsätzliche menschliche Eigenschaften hervorgerufen werden. Der Versuch den Bediener zu entlasten, kann zu einer Unterforderung führen. Diese führt wiederum zu einer Abnahme der Aufmerksamkeit, wodurch die Wahrscheinlichkeit von Fehlentscheidungen erhöht wird. Hierdurch wird deutlich, dass die Entwicklung des technischen Systems ebenso eine Rolle spielt, wie die Anpassung des technischen Systems an den Menschen [VDI89], [Gei90].

Dass der Mensch einen vermeintlich hohen Anteil am Unfallgeschehen hat, zeigen etliche Unfallstatistiken. Je nach Quelle werden bis zu 70 Prozent der Vor- oder Unfälle menschlichen Fehlern zugeschrieben [And99], [Dhi07]. Eine Auswertung von Unfällen¹³ mit Reisenschäden im Bereich der Deutschen Bahn AG (DB AG) im Zeitraum von 1994 bis 2001 ergibt, dass der technisch gesicherte Eisenbahnbetrieb um einen Faktor 100 sicherer ist als der menschliche Regelbetrieb und um einen Faktor 10.000 sicherer als der Betrieb in der Rückfallebene [Bra05]. Das Bundesamt für Verkehr schreibt in seinem Sicherheitsbericht aus dem Jahr 2014, dass die meisten Unfälle auf menschliches Fehlverhalten zurückzuführen sind. Hierzu zählen aber nicht nur Fehler des Personals, sondern auch die Unaufmerksamkeit von Reisenden und die Gedankenlosigkeit von Drittpersonen [BAV15]. Um die Ursachen hierfür besser verstehen zu können, wird im nachfolgenden Kapitel näher auf die Entstehung menschlicher Fehler eingegangen.

¹³ Eine umfassende Übersicht über Eisenbahnunfälle bei der DB AG liefert PREUß [Pre08].

2.1.2.1 Ursachen für die Entstehung menschliche Fehler

Richtiges Handeln und eingetretene Fehler sind nach *REASON* die zwei Seiten derselben Medaille [Rea90]. Der Autor beschreibt eine kognitive Bilanz und sieht menschliche Fehlhandlungen als Tribut, den der Mensch für seine bemerkenswerte Fähigkeit, die Welt zu modellieren und zu vereinfachen, zu zahlen hat. Mit anderen Worten, geht mit jeder korrekten Ausführung auf der Aktivseite ein entsprechender systematischer Fehler als Schuldposten auf der Passivseite einher. Diese bildhafte Darstellung ist eher eine metaphorische Beschreibung als eine Erklärung und somit ungeeignet für die Beschreibung der Entstehung menschlicher Fehler. Ursachen für menschliche Fehler sind vielschichtig und im operativen und/oder organisatorischen Umfeld zu suchen.

Ein großes Problem im heutigen Arbeitsalltag sind monotone Arbeitsbedingungen. Hierdurch schwindet die Aufmerksamkeit und die Fehleranfälligkeit des Menschen wird erhöht (Kapitel 2.1.1.2.1). Ein weiterer Punkt, der die Fehlerwahrscheinlichkeit negativ beeinflusst, ist der größer werdende Anteil an komplexen Systemen [And09] und das immer umfassender werdende Regelwerk. Komplexität bedeutet in diesem Kontext nicht, dass die Anlage aus vielen Teilen besteht. Sie wird vielmehr als Undurchschaubarkeit definiert und begünstigt die Unvorhersehbarkeit von ungewollten Zuständen. Der Fehlerwirkweg ist unvorhersehbar und die Fehlerabwehrbarrieren¹⁴ sind abhängig voneinander.

Das Problem daran ist, dass die Wirksamkeit der Fehlerabwehrbarrieren durch diese Abhängigkeit reduziert wird. Das Umgehen einer Barriere kann zum gleichzeitigen Aushebeln einer anderen führen [Elm01] und einen gefährlichen Zustand begünstigen. Hier liegt ein erster Ansatzpunkt zur Verringerung menschlicher Fehler in technischen Systemen. Wären die Barrieren klar ersichtlich und unabhängig, könnte ein Großteil der Fehler abgeblockt werden. Gerade in der Rückfallebene ist demnach ein verständliches Regelwerk notwendig, um ein akzeptables Sicherheitsniveau aufrechtzuerhalten.

2.1.2.2 Umgang mit Komplexität der Technik und Abhängigkeit der Faktoren

Um die Themen Komplexität und Abhängigkeit nochmals aufzugreifen, werden nachfolgend vier Faktoren genannt, die die Sicherheit und Zuverlässigkeit im Eisenbahnwesen maßgeblich beeinflussen [And99], [Elm01]:

- die Sicherungstechnik,
- ein klar formuliertes Regelwerk,
- qualifiziertes Personal und
- ein Managementsystem mit Berichtswesen.

Bereits auf den ersten Blick fällt auf, dass die genannten Faktoren voneinander abhängig sind. Ein Ausfall der Sicherungstechnik führt zu einem energieärmeren Zustand, d. h. zu ge-

¹⁴ siehe z. B. das Schweizer-Käse-Modell [Rea90]

ringeren Geschwindigkeiten oder zum Stillstand. Das System fällt auf die Rückfallebene zurück und der Fdl übernimmt die Verantwortung. Kommt es infolgedessen zu einer kritischen Situation oder einem Unfall, ist es einfach, von menschlichem Versagen zu sprechen. Diese unüberlegte Schuldzuweisung ist aber nicht zielführend. In der Rückfallebene stützt sich der Fdl, als qualifiziertes Personal, auf sein vorhandenes Wissen und das betriebliche Regelwerk. Das Regelwerk beschreibt ihm das genaue Vorgehen und die regelkonforme Anwendung der Hilfshandlungen. Ist es z. B. kompliziert oder widersprüchlich verfasst, wird dem Fdl die korrekte Ausführung erschwert und Fehlentscheidungen begünstigt. Hinzu kommt, dass die Fehlermöglichkeiten in der Rückfallebene mit der Anzahl der Beteiligten und den stets größer werdenden Zuständigkeitsbereichen wachsen. Beispielsweise sind die großen Zuständigkeitsbereiche der Grund dafür, dass sich die Anzahl der Hilfshandlungen nicht verringert, sondern erhöht. Zum Schluss fehlt noch das Managementsystem. Dieses beinhaltet die Elemente Sicherheitsmanagement, Qualitätsmanagement, Arbeits- und Brandschutz sowie Umweltschutz. Das System stellt die Umsetzung der geltenden rechtlichen Bestimmungen und der unternehmerischen Vorgaben sicher und befasst sich zudem mit der kontinuierlichen Verbesserung von Prozessen. Es werden Vor- und Unfälle dokumentiert und analysiert, um Lerneffekte anzustoßen, damit z. B. das Regelwerk angepasst und die Systemgestaltung auf menschliche Belange ausgerichtet werden kann.

2.1.2.3 Exkurs Klassifizierung menschlicher Fehler

Um die Ursachen des Fehlverhaltens besser beschreiben zu können, werden menschliche Fehler nach ihrem Eintreten und ihrer Entstehung unterschieden. Bevor mit der Klassifizierung begonnen werden kann, wird zunächst erläutert, was unter dem Begriff *menschlicher Fehler* zu verstehen ist. Eine systembezogenen Definition liefert die DIN EN 50129 ([DIN03], S. 12):

„Menschliche Fehler sind menschliche Handlungen, die zu einem ungewollten Verhalten des Systems oder zu einer Fehlfunktion führen können.“

Die Fülle an Klassifizierungsansätzen menschlicher Fehler wird in zwei große Bereiche aufgeteilt: auftretensorientierte und ursachenorientierte Klassifizierung. Diese Aufteilung ergibt sich durch unterschiedliche Betrachtungsstandpunkte, da eine hundertprozentige Abtrennung nicht immer möglich ist [Joh93]. Die erste Klassifizierung bezieht sich auf einen mehr verhaltenspsychologischen Betrachtungsstandpunkt, bei dem die Fehler entsprechend ihres Auftretens (z. B. sporadisch, zufällig, systematisch) unterschieden werden. Bei der zweiten Gruppe wird zusätzlich die Frage nach dem *Warum?* gestellt. Hier liegt der Fokus auf den Prozessen des Handelns, der Motivation und der Informationsverarbeitung [Hof08]. Zudem wird das Auftreten von Systemfehlern untersucht, um durch bessere Systemgestaltung menschliche Fehler zu verhindern [Bub92] und Lerneffekte anzuregen [GiTi06], [Hof08].

Im Folgenden werden zwei Ansätze genauer beschrieben – das Drei-Ebenen-Modell von *RASMUSSEN* und die übergreifende Klassifizierung von *REASON*.¹⁵ Ersterer gehört zur ursachenorientierten Klassifizierung, wohingegen *REASON* einen systemübergreifenden Ansatz verfolgt und die o. g. Klassifizierungsansätze kombiniert.¹⁶ Auf die Beschreibung eines Ansatzes aus der Kategorie der auftretensorientierten Klassifizierung wird an dieser Stelle verzichtet. Diese Art der Klassifizierung bezieht sich fast nur auf systemnotwendige Tätigkeiten des Bedieners, in Form von motorischen Handlungen, und berücksichtigt kaum kognitive Tätigkeiten, wie sie bei Problemlösungs- und Entscheidungsaufgaben notwendig sind [Bub92].

2.1.2.3.1 Drei-Ebenen-Modell menschlichen Verhaltens nach *RASMUSSEN*

Ursprünglich war das Modell dazu gedacht, menschliches Fehlverhalten in kerntechnischen Anlagen zu erfassen und zu klassifizieren. Aus heutiger Sicht reicht die Bedeutung und Anwendbarkeit jedoch weit über diesen Bereich hinaus und das Modell dient als Standard für die Einteilung menschlicher Fehler in vielen Industriezweigen. *RASMUSSEN* konstatiert, dass menschliche Fehler von einer Vielzahl an Faktoren abhängig sind und, je nach Wissensstand der Person, verschiedene Prozessebenen durchlaufen werden (Abbildung 2-4). Diese unterscheiden sich in der Art der Fehler und der kognitiven Abläufe voneinander. So ist zum Beispiel das Verhalten auf der wissensbasierten Ebene von einer sequenziellen Aufgabenabarbeitung geprägt, wohingegen sich die fertigkeitsbasierte Ebene durch eine simultane Aufgabenbewältigung auszeichnet [Ras86], [Bub92]. Nachfolgend werden die drei Ebenen genauer beschrieben.

Fertigkeitsbasiertes Verhalten

Auf dieser Ebene wird das Verhalten bei oft ausgeübten Aufgaben, für die ein geringer Grad bewusster Denktätigkeit erforderlich ist, beschrieben [VDI15]. Es sind keine wissentliche Aufmerksamkeit und Steuerung mehr notwendig. Dazu gehören Routineaufgaben und wiederkehrende Handlungsabläufe wie z. B. die Bedienung der Sicherheitsfahrschaltung (Sifa)¹⁷ durch den Tf [Pet80]. Aufgaben auf dieser Ebene lassen einen Spielraum für Nebentätigkeiten.

Fehler auf dieser Ebene entstehen in der Regel durch Unaufmerksamkeit oder Überaufmerksamkeit. Das bedeutet, dass Handlungen falsch oder zur falschen Zeit ausgeführt werden.

¹⁵ Thema der vorliegenden Arbeit ist die Untersuchung des Einflusses der farblichen Gestaltung der Bedienoberfläche auf den Fahrdienstleiter. Die Entwicklung einer neuen Klassifizierung bzw. Quantifizierung menschlicher Zuverlässigkeit bezogen auf die Systemsicherheit ist nicht Bestandteil. Deshalb fand bei der Auswahl der vorgestellten Klassifizierungen für menschliche Fehler keine genauere Betrachtung von eventuell vorhandenen Schwächen oder Untersuchungen bezogen auf die Anwendungstauglichkeit des Modells statt. Es handelt sich ausschließlich um die populärsten Modelle, welche bis heute als Grundlage für viele Forschungsansätze dienen.

¹⁶ Auf der Grundlage des Drei-Ebenen-Modells entwickelt *RASMUSSEN* später einen Multi-Aspekt Ansatz, welcher ebenfalls eine Kombination aus beiden Ansätzen darstellt [Ras87].

¹⁷ für eine Funktionsbeschreibung der Sifa siehe Kapitel 2.1.3.2.1

Regelbasiertes Verhalten

Auf der regelbasierten Ebene werden weniger vertraute Aufgaben gelöst. Das Verhalten resultiert aus dem Vergleich der Informationen mit vertrauten Mustern und Regeln, welche anschließend zu einer neuen Verhaltensweise kombiniert werden [VDI15].

Fehler entstehen, wenn die Situation aufgrund fehlender Informationen falsch eingeschätzt wird. Das führt dazu, dass die bekannte Verhaltensregel falsch oder eine falsche Verhaltensregel angewendet wird.

Wissensbasiertes Verhalten

Der Umgang mit meist neuartigen und ungeübten Aufgaben wird auf der wissensbasierten Ebene beschrieben. Es ist ein hoher Grad an bewusster Denktätigkeit erforderlich, da vertraute Muster und Regeln nicht direkt angewendet werden können. Handlungspläne müssen komplett neu erstellt werden [VDI15]. In dieser Ebene ist äußerst viel Transferleistung erforderlich, was den Menschen besonders anfällig für Fehler macht.

Fehler in dieser Ebene entstehen dadurch, dass fehlendes oder falsches Wissen auf eine gänzlich unbekannte Situation angewendet wird.

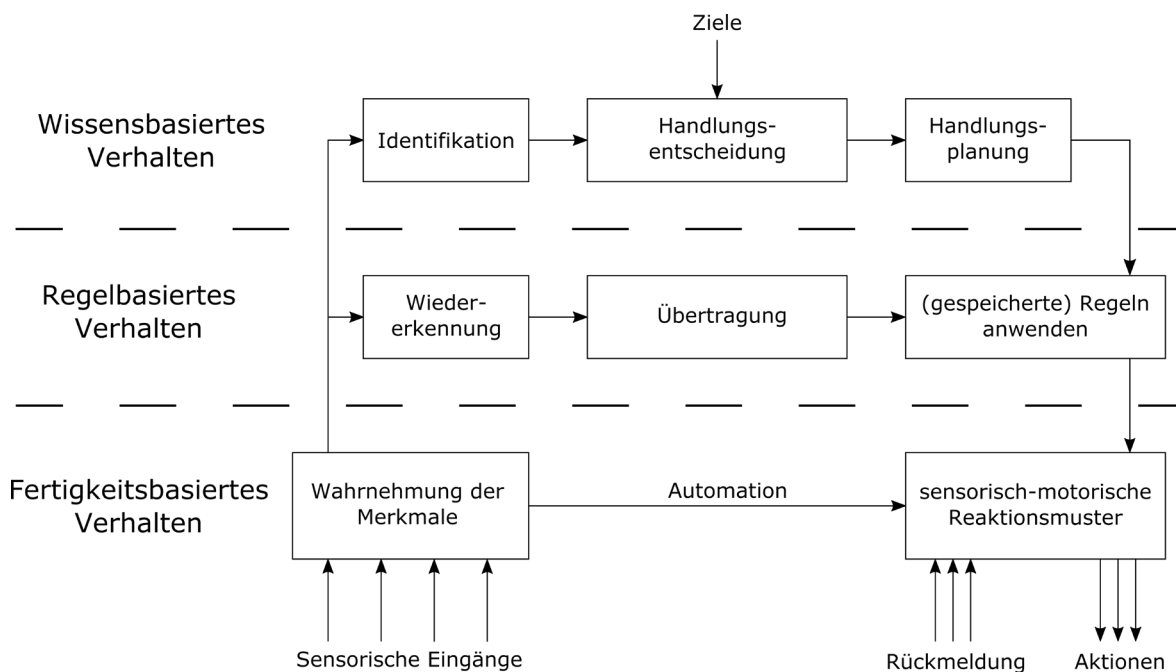


Abbildung 2-4: Schematische Darstellung der kognitiven Ebenen des menschlichen Verhaltens nach RASMUSSEN [Ras86] (Darstellung entnommen aus [And09], vereinfacht durch den Verfasser)

Das Modell ist so zu verstehen, dass je nach Aufgabe alle drei Ebenen durchlaufen werden (müssen). Handelt es sich um eine Routineaufgabe in einer bekannten Umgebung, wird diese automatisch ausgeführt, was wiederum direkt zum Zielzustand führt (fertigkeitsbasiertes Verhalten). Ist das nicht möglich, muss die Situation analysiert werden. Sind Muster vertraut, werden diese mit gespeicherten Verhaltensregeln assoziiert. Ist das erfolgreich, führt dies

zum Zielzustand (regelbasiertes Verhalten). Werden keine vertrauten Muster erkannt, müssen abstrakte Beziehungen zwischen der Struktur und der Funktion der Aufgabe analysiert werden. Hierbei wird auch die Zielstellung bei der Entscheidungsfindung mitbetrachtet (Transferleistung). Anschließend wird die Handlung geplant (wissensbasiertes Verhalten). Können jetzt gespeicherte Verhaltensregeln angewendet werden, führt deren Anwendung zum Zielzustand. Ist das nicht der Fall, werden neue Verhaltensregeln formuliert und anschließend angewendet.

2.1.2.3.2 Einteilung menschlicher Fehler nach *REASON*

Aufgrund einer Häufung spektakulärer Unfälle, wie z. B. der Nuklearkatastrophe von Tschernobyl, rückten die Ausgangsbedingungen von Unfällen in den Fokus. Es entwickelte sich eine systemorientierte Sichtweise, da neben einer falschen Handlung auch eine ungünstige Verkettung von Handlungen zu Unfällen führen kann. Ein häufig genannter Vertreter dieser Bewegung ist die Klassifizierung von *REASON*.

Dieser Ansatz vereint die auftretensorientierte und ursachenorientierte Sicht [Rea90], [Rea94]. Der Autor kombiniert die *Activation-Trigger-Theory* von *NORMAN* [Nor81] und das Drei-Ebenen-Modell von *RASMUSSEN* [Ras86] in seinem Modell (Abbildung 2-5). Im Gegensatz zu den übrigen Modellen klassifiziert der Autor nicht nur Fehler, sondern auch Handlungen, die zu einer Verringerung der Sicherheit führen können [Hof08].

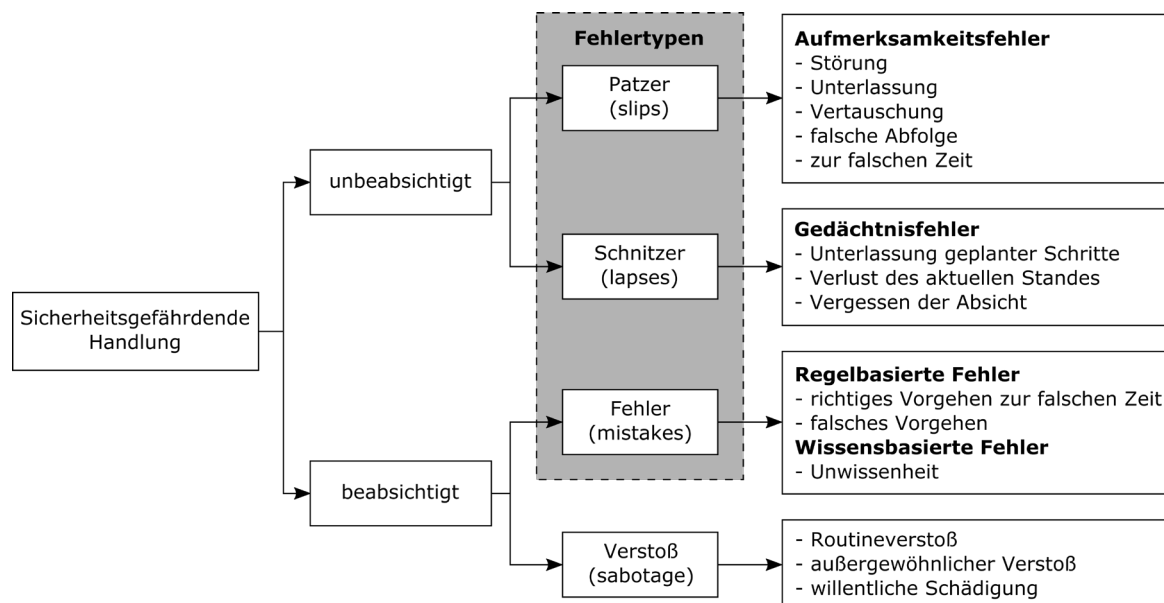


Abbildung 2-5: Einteilung sicherheitsgefährdender Handlungen nach *REASON* [Rea90] (Darstellung entnommen aus [And09] und geringfügig verändert)

REASON unterscheidet die Fehler zum einen nach der Ebene der Handlungskontrolle auf der sie vorkommen und zum anderen danach, ob sie beabsichtigt sind oder nicht. Wird eine Handlung bewusst durchgeführt und erreicht dennoch nicht das angestrebte Ziel, kann dies zwei Gründe haben: der Fehler liegt in der Ausführung oder in der Planung. Der Misserfolg tritt ein, weil die Handlung z. B. zur falschen Zeit oder abweichend vom geplanten Ablauf

durchgeführt wird. Der Autor spricht in diesen Fällen von Patzern (*slips*) und Schnitzern (*lapses*). Erfolgt die Durchführung nach Plan, führt aber nicht zum gewünschten Ergebnis, spricht der Autor von (Denk-)Fehlern (*mistakes*). Die beabsichtigte Folge tritt nicht ein, weil der Plan auf schlechten oder falschen Annahmen beruht oder nicht angemessen ist, d. h. die Situation wird falsch eingeschätzt. Die letzte Unterscheidung befasst sich mit Verstößen (*sabotage*). Hierzu zählen bewusste Abweichungen vom Regelwerk (z. B. durch Selbstüberschätzung) oder die willentliche Schädigung.

2.1.3 Exkurs Human Factors im Eisenbahnwesen

Die vorangegangenen Erklärungen zeigen, dass Ursachen für menschliche Fehler mitunter in der Systemgestaltung und den Arbeitsbedingungen gesucht werden können. Im nächsten Schritt wird ein Forschungsbereich vorgestellt, welcher sich mit der Analyse von Schwachstellen im Arbeitsumfeld des Eisenbahnwesens auseinandersetzt. Hierbei werden auch menschliche Einflussfaktoren ins Kalkül einbezogen.

2.1.3.1 Abgrenzung *human factors* und *ergonomics*

Ist die Rede von Human Factors, werden hierunter oftmals menschliche Einflussfaktoren innerhalb eines technischen Systems verstanden [BHL08]. Diese leichtfertige Verwendung greift zu kurz, verschärft die Wahrnehmung des Menschen als Risikofaktor und spiegelt den Forschungsbereich nicht in seiner Vollständigkeit wider. Die Human-Factors-Forschung beschäftigt sich mit der Mehrdimensionalität und Komplexität der psychologischen, sozialen und technischen Einflussfaktoren [StHo14]. Schon sehr früh wird die Notwendigkeit eines interdisziplinären Forschungsansatzes erkannt, um Probleme im Zusammenspiel zwischen Mensch und Maschine lösen zu können.¹⁸ Deshalb steht im Vordergrund der Human-Factors-Forschung die Verknüpfung menschlicher und technischer Einflussfaktoren. Eine weitverbreitete und anerkannte Definition stammt von der *International Ergonomics Association* (IEA), welche ebenfalls von der *Human Factors and Ergonomics Society* übernommen wurde [IEA00], [HFES00]:

„Ergonomics (or human factors) is the scientific discipline concerned with the understanding of interactions among humans and other elements of a system, and the profession that applies theory, principles, data and methods to design in order to optimize human well-being and overall system performance.

Practitioners of ergonomics and ergonomists contribute to the design and evaluation of tasks, jobs, products, environments and systems in order to make them compatible with the needs, abilities and limitations of people.

Ergonomics helps harmonize things that interact with people in terms of people's needs, abilities and limitations.”

¹⁸ siehe hierzu exemplarisch [Lom64]

Die Definition zeigt, dass die Begriffe *human factors* und *ergonomics* im englischsprachigen Raum synonym verwendet werden.¹⁹ Diese Analogie lässt sich jedoch nicht auf die deutsche Sprache übertragen. Ergonomie befasst sich in Deutschland in erster Linie mit körperlichen und geistigen Aspekten und weniger mit kognitiven Prozessen und Motiven. Forschungsbestrebungen im Bereich *ergonomics* werden im deutschsprachigen Raum in das Gebiet der Arbeitswissenschaften eingeordnet, wohingegen die Erforschung von Denkprozessen Aufgabe der angewandten Psychologie ist.

Wie eingangs erwähnt, dient eine Vielzahl von Wissenschaften²⁰ als Grundlage für die Beschreibung der Interaktion des Menschen mit technischen Systemen. Abbildung 2-6 verdeutlicht den Zusammenhang und zeigt die beteiligten Wissenschaften, welche die Grundlagen für die vorliegende Arbeit schaffen. Die Schnittmenge kennzeichnet den Bereich der Human-Factors-Forschung. Für nähere Informationen zu den Forschungsbestrebungen der Ergonomie und der angewandten Psychologie wird an dieser Stelle auf Kapitel 2.2.1.3 verwiesen.

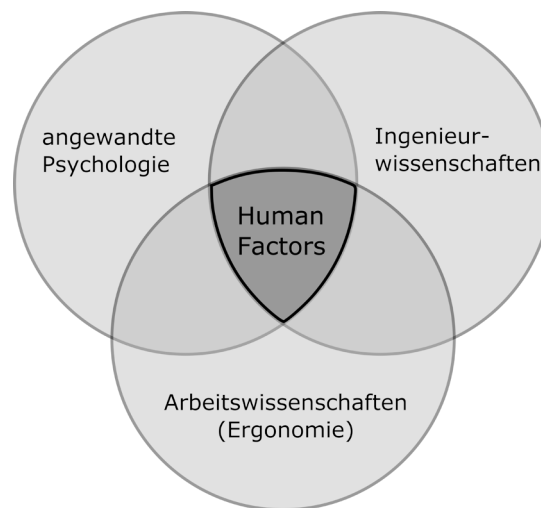


Abbildung 2-6: Basiswissenschaften der Human-Factors-Forschung

2.1.3.2 Rail human factors

Seit Anfang des 21. Jahrhunderts finden vermehrt Studien statt, welche sich mit menschlichen Einflussfaktoren und psychologischen Prozessen im Eisenbahnwesen befassen. Diese Forschungsbestrebungen werden unter dem Begriff *rail human factors*²¹ zusammengefasst. Im Interesse der *rail human factors* stehen die Folgen, die sich aus den geänderten Arbeitsbedingungen für die Mitarbeiter im Eisenbahnwesen ergeben. Eine Vorreiterrolle in Sachen *rail human factors* nimmt Großbritannien ein. Im *Rail Safety and Standards Board* (RSSB) fanden bis dato die weitreichendsten Studien statt. Eine umfangreiche Sammlung der Er-

¹⁹ siehe hierzu auch [Kar12]

²⁰ In der vorliegenden Arbeit wird bewusst zwischen den Begriffen *Wissenschaft* und *Disziplin* unterschieden. Wissenschaft beschreibt die Gesamtheit der Erkenntnisse einzelner Gebiete und befasst sich mit der systematischen Vermehrung, Darstellung und Begründung eben dieser. Eine Disziplin umfasst ausschließlich die Ausbildung der Erkenntnisse. Statt des Begriffs *Disziplin* kann auch der Begriff *Fach* verwendet werden [Zim06].

²¹ Eine Definition der *Human factors* im Eisenbahnwesen findet sich in [HJL08].

gebnisse liefern die *Proceedings* der letzten fünf *International Conference on Rail Human Factors* [WNC+05], [WNC+07], [WMC+12], [WMD+13] und [WMD+15]. Des Weiteren haben *WILSON UND NORRIS* einen Artikel über die aufkeimende Wissenschaft geschrieben [WiNo05]. Die Autoren fassen die englischsprachige Literatur zusammen und geben in ihrem Werk einen Überblick über die aktuelle Forschung sowie zukünftige Vorhaben.

Im Folgenden werden weitere Erkenntnisse der *rail human factors* aufgeführt. Der Fokus liegt dabei auf der Analyse leistungsbeeinflussender Faktoren für die Berufsgruppe der Triebfahrzeugführer und der Fahrdienstleiter.

2.1.3.2.1 Schwerpunkt Triebfahrzeugführer

Grundsätzlich werden für die Berufsgruppe der Tf leistungsbeeinflussende Faktoren wie Aufmerksamkeit, Wahrnehmung und Ablenkung genannt. Zwei wichtige Punkte sind hierbei die Entstehung und die Folgen von Monotonie²². Die Gründe für das Auftreten von Monotonie sind vielfältig. Sie reichen von personenbedingten Ursachen, wie ein gestörter Tag-Nacht-Rhythmus durch Schichtarbeit²³, über mangelhafte ergonomische Gestaltung des Arbeitsplatzes, bis hin zur Unterforderung durch z. B. reizarme Umgebung oder mangelnde Abwechslung während der Fahrt. Es kommt vor allem bei langsamen Geschwindigkeiten und in der Dunkelheit zur psychischen Ermüdung [RLM02], [BiSc10]. Als Folge nimmt die anhaltende Aufmerksamkeit (Vigilanz) ab. Der Begriff *Vigilanz* beschreibt den Zustand oder den Grad der Bereitschaft, zufällige, unvorhersehbare und mit geringer Wahrscheinlichkeit auftretende kritische Signale, unter monotonen Bedingungen zu entdecken. *MACKWORTH* hat als Erster dieses Phänomen systematisch beschrieben [Mac57]. Er zeigt in einem Experiment, dass der Vigilanzabfall degressiv verläuft und in der ersten halben Stunde am größten ist [Kra85]. Vigilanzminderung wird auch im Eisenbahnwesen immer häufiger als eigentliche Unfallursache gesehen [WiSt83]. Auslöser hierfür ist, dass der Mensch seine eigene Wachheit eher schlecht einschätzen kann. In Versuchsreihen wurde herausgefunden, dass das subjektive und das psychologische Maß der Müdigkeit nur geringfügig miteinander korrelieren [MSF09].

Aus diesem Grund werden unterschiedliche Einrichtungen zur Kontrolle des Tf im Triebfahrzeug installiert. Als bekannteste Beispiele sind die Sicherheitsfahrschaltung (Sifa) und die Indusi-Wachsamkeitstaste zu nennen.²⁴ Interessanterweise ergaben Untersuchungen in der Vergangenheit, dass die aktuell im Einsatz befindlichen Einrichtungen gar nicht oder nur un-

²² Zustand herabgesetzter psychischer Aktivität [Nit71]

²³ Da die Müdigkeit einen Einfluss auf fehlerhafte Bedienhandlungen hat, sind gewisse Uhrzeiten für Unachtsamkeit prädestiniert. Eine Auswertung der Anzahl an Zwangsbremssungen und akustischen Sifa-Handlungsaufforderungen hat ergeben, dass gegen 3 Uhr nachts und 15 Uhr am Nachmittag sowohl die meisten Zwangsbremssungen als auch Sifa-Handlungsaufforderungen auftreten. Dieser Effekt verstärkt sich zusätzlich, wenn die kritischen Zeiten am Ende einer Schicht liegen [HRR74].

²⁴ Mit der Wachsamkeitstaste quittiert der Tf die 1000 Hz-Beeinflussung der punktförmigen Zugbeeinflussung, die ein Halt zeigendes Signal ankündigt. Bei der Sifa betätigt der Tf ein Pedal, wobei der Druck alle 30 s gelöst werden muss. Erfolgt dies nicht, warnt das System den Tf zunächst durch eine optische Anzeige. Anschließend ertönt ein zusätzliches akustisches Warnsignal. Reagiert der Tf wiederholt nicht, wird der Zug zwangsgebremst.

zureichend in der Lage sind, die Aufmerksamkeit der Tf zu unterstützen, respektive zu steigern, weil systemimmanente Gewöhnungseffekte die Wirkung zunichtemachen [Pet80], [WPM05]. So wird die Sifa noch ordnungsgemäß betätigt, obwohl im Laufe der Zeit die Dauer der Fußpedalbetätigung, der Kraftaufwand für die Bedienung und die Reaktionszeit durch die verringerte Aufmerksamkeit abnehmen [CCM+93].

Um diesem Phänomen entgegenzuwirken, nimmt der Automatisierungsgrad des Tf-Arbeitsplatzes stetig zu und Themen wie der vollautomatisierte Zugbetrieb rücken in den Vordergrund [S2R16]. Die linienförmige Zugbeeinflussung (LZB) und das *European Train Control System* (ETCS) ermöglichen bereits heute eine hohe Systemautonomie. Unklar dabei ist jedoch, ob durch diese Systeme wirklich eine Verbesserung der Arbeitsbedingungen entsteht. Hier liegt die Vermutung nahe, dass die Unterstützung eher die Unterforderung fördert und die Vigilanzminderung begünstigt. Um die Aufmerksamkeit des Tf positiv zu beeinflussen, werden die freiwerdenden Kapazitäten durch neue Aufgaben ausgefüllt. Mit der Informations- und Kommunikationsplattform Rail-in-Motion (RiM), werden dem Tf seit geraumer Zeit situationsgerecht weiterführende Informationen, wie z. B. Fahrempfehlungen für pünktliches und energieeffizientes Fahren, auf einem Tablet zur Verfügung gestellt [Zöl15]. Dieses Vorgehen ist kritisch zu hinterfragen, denn nicht jeder Mitarbeiter ist der ansteigenden Informationsflut gewachsen oder geht mit dieser Möglichkeit verantwortungsvoll um.

Die Ergebnisse für die Berufsgruppe des Tf zeigen, dass sich technische Einrichtungen nicht immer eignen, um die Aufmerksamkeit zu überwachen und/oder zu erhöhen. Soziale Komponenten, in Form eines sozialverträglichen Schichtplans, sind ebenso wichtig wie eine intelligente technische Unterstützung. Zudem sind Änderungen am Arbeitsplatz mit allen Beteiligten abzustimmen und freie Kapazitäten sind sinnvoll auszufüllen, um die Zuverlässigkeit des Gesamtsystems Bahn zu verbessern.

2.1.3.2.2 Schwerpunkt Fahrdienstleiter

Analysen für den Fahrdienstleiterarbeitsplatz haben die (mentale) Arbeitsbelastung sowie die Gestaltung der Anzeigen als leistungsbestimmende Faktoren identifiziert [HMD+82]. Die Schwerpunkte der Arbeitsbeanspruchung liegen im Bereich der Informationsaufnahme und der Informationsverarbeitung, aber auch bei den arbeitsnotwendigen Kontakten – im Sinne von Kommunikation. Hinzukommen die dauernde Aufmerksamkeit und Konzentration sowie das vermehrte Unterbrechen von Handlungen und Denkprozessen [ScSc88]. Weitere Faktoren sind der Aufwand für Hilfshandlungen, die Stellbereichsgröße und die damit verbundene Zahl der abzufertigenden Züge [Hey07] sowie das Auftreten von Störungen [MDS+91].

Ein Großteil der Faktoren wird im modernen Stellwerk maßgeblich durch den Bedienplatz bzw. die Bedienoberfläche beeinflusst. Hier herrscht jedoch großer Forschungsbedarf. Ein Blick auf heutige Bedienoberflächen zeigt, dass hier definitiv Nachholbedarf besteht. Die farbliche Gestaltung und die grafische Darstellung der Anzeigen haben sich in den letzten 30 Jahren kaum verändert [Pac16].

Vor allem die farbliche Gestaltung der Stellwerksbedienoberfläche wird in Zukunft einen höheren Stellwert einnehmen müssen. Erkenntnisse aus anderen Industriezweigen zeigen, dass sich der gezielte Einsatz von Farben, positiv auf die Informationsverarbeitung und den Vigilanzabfall auswirken kann. Analysen zeigen, dass eine ansprechende farbliche Gestaltung der elektronischen Anzeigen den Bediener stimuliert [PfWi86], [PfWi87]. Zudem spielt die Farbe bei der Objektwahrnehmung eine bedeutsame Rolle, weil sie die Leistung der visuellen Informationsverarbeitung unterstützt [Chr75], [DaOs88]. Auch die subjektive Wahrnehmung der Bedienungssicherheit wird durch eine farblich-ästhetische Darstellung erhöht [Wan93], wodurch die Arbeitsfreude und die Leistungsbereitschaft gesteigert werden können [MoTh10].

2.1.4 Zusammenfassung Kapitel 2.1

Angesichts der geschilderten Sachverhalte wird deutlich, dass der Mensch nicht allein für einen Vor- bzw. Unfall verantwortlich gemacht werden kann und der verwendete Ausdruck *menschliches Versagen* oft keine aussagekräftige Begründung ist. Die begrenzte Informationsaufnahme und die Fehleranfälligkeit in der Entscheidungsfindung sowie die manchmal träge Reaktion sind systemimmanente Eigenschaften des Menschen. Demgegenüber stehen Fähigkeiten wie Kreativität, Fehler selbst zu erkennen und auf unvorhergesehene Situation zielorientiert reagieren zu können. Eine Maschine fällt aus und muss repariert oder ausgetauscht werden. Ein Mensch hingegen hat die Möglichkeit seine eigene Leistung zu überwachen, auf Fehler zu reagieren und diese zu korrigieren bevor sie die Systemleistung beeinflussen. Aus diesen Gründen bleibt der Mensch ein gleichberechtigter Teil des Systems. Nicht zuletzt bilden Handlungsfehler eine Möglichkeit Lernprozesse anzuregen. Diese dienen dann als Grundlage für individuelles und organisationales Lernen.

Die fehlerfokussierte Sichtweise hat den Automatisierungsprozess vorangetrieben. Der Mensch wurde in der Vergangenheit – und wird es stellenweise auch heute immer noch – als schwächstes Glied der Kette gesehen. Deshalb soll er nach und nach ersetzt werden. Dies führt jedoch zu einer teilweise paradoxen Rolle des Menschen im Gesamtsystem (*Ironies of Automation*) und einer Reihe neuartiger Fehlerarten. Sensomotorische Fähigkeiten verlieren an Bedeutung, wohingegen das Thema der Vigilanzminderung und die kognitiven Prozesse während der Entscheidungsfindung mehr und mehr an Bedeutung gewinnen. Hinzu kommt, dass die hochautomatisierte Anlage den Fdl im Regelbetrieb ausschließt. Er bleibt *out-of-the-loop*. Da er eine passive Rolle einnimmt, ist er im Störfall möglicherweise nicht ausreichend auf die Situation eingestellt und kann somit die Zusammenhänge und das Problem nicht schnell genug erfassen. Aus diesem Grund muss eine Veränderung zu einem sichereren und zuverlässigeren Verkehrssystem Bahn auf verschiedenen Ebenen stattfinden. Eine höhere Systemzuverlässigkeit beschränkt sich nicht allein auf verbesserte technische Komponenten. Sie beinhaltet zusätzlich die Abschaffung organisatorischer Unklarheiten und die Optimierung hinsichtlich einer reibungsarmen Mensch-Maschine-Interaktion. Besonders die Anwendung bestehender betrieblicher Regelwerke auf neu entwickelte Techniken sollte,

aus Sicht des Verfassers der vorliegenden Arbeit, kritisch hinterfragt werden. Aktuell werden die Möglichkeiten der Technik nicht immer in vollem Umfang genutzt, da die neue Technik eher an das bestehende Regelwerk vergangener Generationen angepasst und nicht über eine Überarbeitung respektive Neuorganisation nachgedacht wird.

Auch wenn durch die modernen automatisierten Systeme eine neue Qualität von Fehlern hervorgebracht wird, ist unbestritten, dass sich die Sicherheit im Eisenbahnwesen um ein Vielfaches erhöht hat. Bezogen auf die kontinuierlich gestiegene Verkehrsleistung, hat die Anzahl der Unfälle deutlich abgenommen. Nichtsdestotrotz kann Automatisierung ohne die Einbeziehung des Menschen, nicht erfolgreich betrieben werden [Lüd15]. Der Ruf nach technischen Lösungen und das Streben nach einem perfekten rechnergestützten Betriebsablauf werden nicht allein zu einer optimalen Unfallfreiheit im Eisenbahnwesen führen. Der Mensch ist für die Funktion des Systems wichtig und auch im Eisenbahnwesen aktuell unabdingbar. Dies gilt es zu verstehen, um die Verbesserung des Gesamtsystems weiter voranzutreiben. Als logische Konsequenz rückt die menschengerechte Gestaltung in den Vordergrund. Ziel ist es, arbeitsförderliche Bedingungen zu schaffen, um dadurch ein Höchstmaß an Sicherheit und Zuverlässigkeit zu erreichen und zu bewahren. Dabei sollten die Anzeigen und die dargestellten Informationen der Überwachungs- und Steuerungstätigkeiten entsprechen und die Kapazitäten der Bediener in ausreichendem Umfang berücksichtigt werden.

2.2 Ansätze und Methoden zur Bewertung menschengerechter Gestaltung

Im Jahre 1965 formuliert *MOORE* ein Gesetz, nach dem sich die Komplexität von integrierten Schaltkreisen etwa alle 18–24 Monate verdoppelt [Moo65]. Ein vergleichbares Muster findet sich Jahre später bei der Entwicklung des Funktionsumfangs von technischen Systemen wieder [IEF13]. Das Problem daran ist, dass die Fähigkeiten des Menschen mit dieser raschen Entwicklung nicht Schritt halten können. Die technische Unterstützung soll vorrangig entlasten, führt im Umkehrschluss aber zu einer verminderten Auseinandersetzung mit dem System und zu einer Unterforderung im Arbeitsalltag. Die damit verbundenen Risiken und Probleme wurden in Kapitel 2.1 beschreiben.

Ziel dieses Kapitels ist es, zu zeigen wie wichtig die verlustfreie Interaktion zwischen Mensch und Maschine ist. Dazu werden die Grundlagen zum Mensch-Maschine-System vermittelt und zwei Wissenschaften vorgestellt, die sich mit der menschengerechten Systemgestaltung befassen. Weiter werden zwei Ansätze näher erläutert, welche sich für die Bewertung der Benutzerfreundlichkeit eignen.

2.2.1 Ansätze der menschengerechten Gestaltung

Das Mooresche Gesetz hat aktuell keine Gültigkeit mehr [Hei15]. Nichtsdestotrotz wächst der Funktionsumfang von technischen Systemen stetig weiter. Vor dem Hintergrund, dass neue Technologien den Bediener nur dann entlasten, wenn die Anzahl der notwendigen Entscheidungen oder die Komplexität der zu berücksichtigenden Informationen abnimmt [ScSc88], wird die Notwendigkeit, neue Technologien auf die Fähigkeiten des Menschen zuzuschneiden, immer wichtiger. Dabei muss die Bedienung einfach und intuitiv sein. Mit anderen Worten spielt die Benutzerfreundlichkeit des Produkts eine entscheidende Rolle.

Die Benutzerfreundlichkeit beschreibt die erlebte Nutzungsqualität während der Interaktion mit dem Produkt. Diese Qualität ist auch am Fahrdienstleiterarbeitsplatz wichtig, wird in modernen Anlagen jedoch immer noch vernachlässigt – was in Kapitel 2.4.3 grob skizziert wird. Um hier ansetzen zu können, wird im folgenden Kapitel das Zusammenspiel zwischen Mensch und technischer Anlage näher beschrieben. Das soll helfen, die Zusammenhänge besser zu verstehen und den Fokus der vorliegenden Arbeit verdeutlichen.

2.2.1.1 Grundlagen des Mensch-Maschine-Systems

Automatisierung und Zentralisierung haben dazu geführt, dass sich Bedienelemente, Anzeigen und das Aufgabenspektrum des Fdl im modernen Stellwerk vollständig gewandelt haben. Gleichwohl haben sich zwei grundsätzliche Charakteristiken kaum verändert. Zum einen die Sicherungsphilosophie und zum anderen die Tatsache, dass der Fdl bei der Aufgabenerfüllung durch eine technische Anlage (Maschine) unterstützt wird.

Im Eisenbahnwesen verwendet der Fdl ein Stellwerk für die Sicherung der Zugfahrten und zur Abwicklung des Betriebs. Dieses zielgerichtete Zusammenwirken und die Verbindung von Aufgabenstellung und -erfüllung werden durch das Mensch-Maschine-System (MMS) definiert. Der generische Begriff *Maschine* beschreibt dabei technische Systeme jeglicher Art [Gei90], [Joh93]. Abbildung 2-7 zeigt das vereinfachte Strukturbild des MMS, welches auf das Eisenbahnwesen übertragen wurde. Als Eingangsvariable dient eine Aufgabe, z. B. das Durchführen einer Zugfahrt. Der Fdl prüft den Fahrweg und stellt die Fahrstraße ein. Dies erfolgt im modernen Stellwerk in der Regel über eine Zweitastenbedienung. Anschließend prüft die Anlage, ob die Fahrstraße möglich ist. Ein Spurplanstellwerk beispielsweise, markiert mittels Such- und Echostrom die benötigten Fahrwegelemente, sucht Flankenschutz, verschließt die Elemente und legt die Fahrstraße fest. Dieser Prozess endet mit der Signalauswahl und das Signal geht auf Fahrt. Der in Abbildung 2-7 dargestellte Rückmeldepfad schließt den Regelkreis und zeigt, dass der Fdl dazu in der Lage ist, Aufgabe und Ergebnis miteinander zu vergleichen, um seinen Arbeitserfolg zu kontrollieren. Alle Einflüsse, die auf den Prozess einwirken, werden als Umwelteinflüsse bezeichnet. Diese Umwelteinflüsse stören die Interaktion, beeinflussen den Fdl, das Stellwerk oder beide und verringern die Zuverlässigkeit der menschlichen Leistung [Bub84], [Bub92].

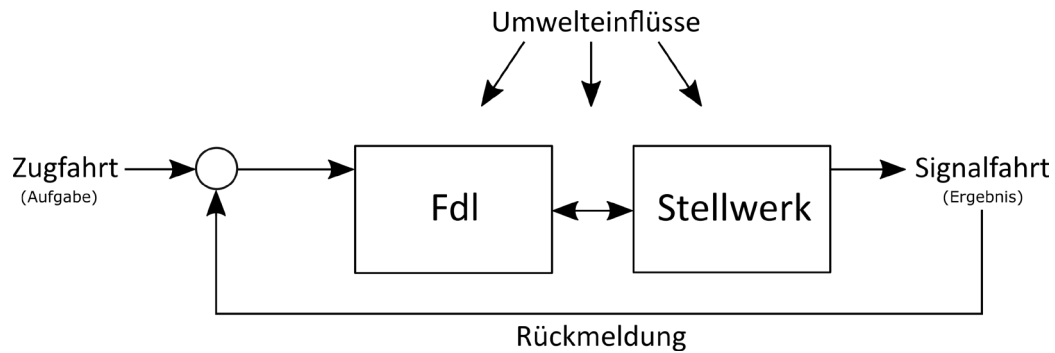


Abbildung 2-7: Das vereinfachte Strukturbild des Mensch-Maschine-Systems (in Anlehnung an [Bub92], durch den Verfasser übertragen auf das Eisenbahnwesen)

2.2.1.2 Informationsaustausch

Eine hochautomatisierte Anlage ist noch kein Garant für ein einwandfrei funktionierendes MMS. Die menschliche Leistungsfähigkeit spielt auch hier eine entscheidende Rolle [Bub77], [Sch75]. Aus diesem Grund ist ein verlust- und reibungsfreier Informationsaustausch zwischen beiden Akteuren (Mensch und Maschine) notwendig, um die Sicherheit und Zuverlässigkeit des Systems zu gewährleisten. Abbildung 2-8 erweitert das vereinfachte Strukturbild des MMS um die Schnittstellen des Informationsaustausches.

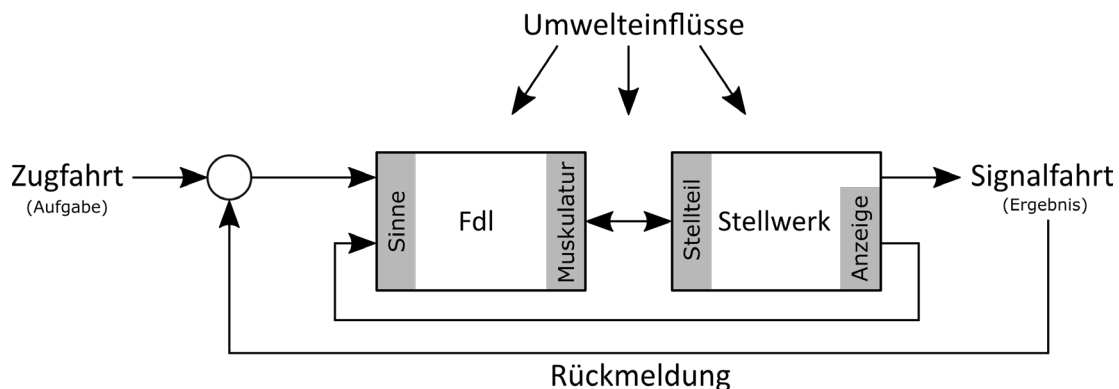


Abbildung 2-8: Das erweiterte Strukturbild des Mensch-Maschine-Systems (in Anlehnung an [Bub92], durch den Verfasser übertragen auf das Eisenbahnwesen)

Über seine Sinnesorgane nimmt der Mensch Informationen aus der Umwelt auf. Diese Informationen werden verarbeitet²⁵, in der motorischen Rinde umgewandelt und schlussendlich über die Muskulatur in eine Handlung umgesetzt. Als Eingangsvariable in dem dargestellten Beispiel dient weiterhin die Aufgabe *Durchführen einer Zugfahrt*. Der Fdl nimmt die Informationen die ihm das Stellwerk mittels Anzeige²⁶ zur Verfügung stellt, wahr, verarbeitet diese und wandelt sie in eine Bedienhandlung um. Auf der technischen Seite bilden Stellteile

²⁵ Die menschliche Informationsverarbeitung hat systemimmanente Eigenschaften, die einerseits sehr hilfreich sind, andererseits aber zu typischen Fehlern führen können. Einen Überblick geben [Wic87], [Mut90] und [Klu06] sowie [WiCa12].

²⁶ Eine Anzeige ist eine Information, die über eine technische Einrichtung den menschlichen Sinnesorganen dargeboten wird [Ber81]. In diesem Fall handelt es sich bei der technischen Einrichtung um den Stelltisch oder die Bedienoberfläche des Bedienplatzes.

oder Bedienelemente die Eingangsseite des Stellwerks. Das Stellwerk verarbeitet die Befehle, führt diese aus und gibt Feedback über die Anzeige. Der Rückmeldepfad dient der Erfolgskontrolle und als Ergebnis geht das Signal auf Fahrt.

2.2.1.3 Erforschung der Mensch-Maschine-Interaktion

Bereits im Jahr 1975 wird ein erster Vorstoß im Eisenbahnwesen gewagt, um Komfort und Zufriedenheit des Bedieners zu erhöhen. Im Mittelpunkt der damaligen Untersuchung steht das Zusammenspiel von Fdl und Stellwerk. *SCHNELL* entwickelt in seiner Dissertation ein Verfahren zur Analyse und Bewertung der Arbeitsschritte eines Fdl im mechanischen Stellwerk [Sch75]. Mit dem entwickelten Verfahren können die Systemelemente ermittelt werden, die die Belastung am Arbeitsplatz im mechanischen Stellwerk maßgeblich beeinflussen. Des Weiteren konstatiert der Autor, dass mit zunehmendem Einsatz von Datenverarbeitungsanlagen die Aufgabenteilung und der Informationsaustausch beider Akteure deutlich mehr zu berücksichtigen sein wird. 40 Jahre später ist die Forderung wichtiger denn je. Der verlust- und reibungsfreie Informationsaustausch trägt maßgeblich zur Verbesserung der Systemleistung und -sicherheit bei. Die Struktur dieses Austausches ist im Grundsatz immer gleich [Bub92] und wird als Mensch-Maschine-Interaktion (MMI) bezeichnet.

Im Folgenden werden zwei Wissenschaften vorgestellt, welche sich im Wesentlichen mit der Erforschung und Verbesserung der MMI auseinandersetzen. Es handelt sich dabei um die Ergonomie, als Teilgebiet der Arbeitswissenschaft, und die Ingenieurpsychologie, als Forschungsbereich der angewandten Psychologie.

2.2.1.3.1 Ergonomie

Für die Übertragung der Information zwischen beiden Akteuren sowie für deren Verarbeitung gelten besondere Gesetzmäßigkeiten. Die Erforschung und Darstellung dieser Wechselwirkungen ist Aufgabe der Ergonomie. Allgemeines Ziel ergonomischer Analysen ist es, die Leistungsfähigkeit des gesamten Arbeitssystems zu optimieren. Dadurch sollen die auf den arbeitenden Menschen einwirkenden Belastungen verringert werden [Sch81a]. Das Erreichen dieses Ziels trägt zur wirtschaftlichen, gesellschaftlichen und umweltbezogenen Nachhaltigkeit bei und erhöht die Akzeptanz der Benutzer. Dabei beinhaltet der ergonomische Gestaltungsansatz biologische, medizinische und psychologische Erkenntnisse und betrachtet u. a. folgende Faktoren [DIN11b]:

- den Zweck des Systems,
- die Eigenschaften der Zielpopulation,
- die Eigenschaften der Aufgabe,
- Beschränkungen (z. B. wirtschaftliche oder gesetzliche Rahmenbedingungen),
- die Arbeitsumgebung und
- den Lebenszyklus.

Um eine optimierte Leistung des Systems zu erreichen, müssen alle Faktoren berücksichtigt werden. Die Faktoren *Zielpopulation*, *Aufgabe* und *Arbeitsumgebung* sind bei einem menschengerechten Gestaltungsansatz besonders zu beachten [DIN11b]. Die Zielpopulation und die Aufgabe sind beides Faktoren, die im Eisenbahnwesen durch die Organisation vorgegeben werden. Im Stellwerk besteht die Zielpopulation aus ausgebildeten und geprüften Fahrdienstleitern, deren Aufgabe die sichere Durchführung des Bahnbetriebs ist. Als Arbeitsumgebung dient der Bedienplatz. Dieser setzt sich aus der Bedienoberfläche und den Bedienelementen zusammen. Er bildet die Benutzerschnittstelle²⁷ zwischen Fdl und Stellwerk und ermöglicht dem Fdl somit die direkte Interaktion mit dem Stellwerk. Demnach nimmt der Bedienplatz eine Schlüsselrolle für den verlust- und reibungsfreien Informationsaustausch zwischen beiden Akteuren ein. Abbildung 2-9 greift das vereinfachte Strukturbild des MMS (Abbildung 2-7) auf, verdeutlicht die Lage der Benutzerschnittstelle im Stellwerk und zeigt die beteiligten Elemente.

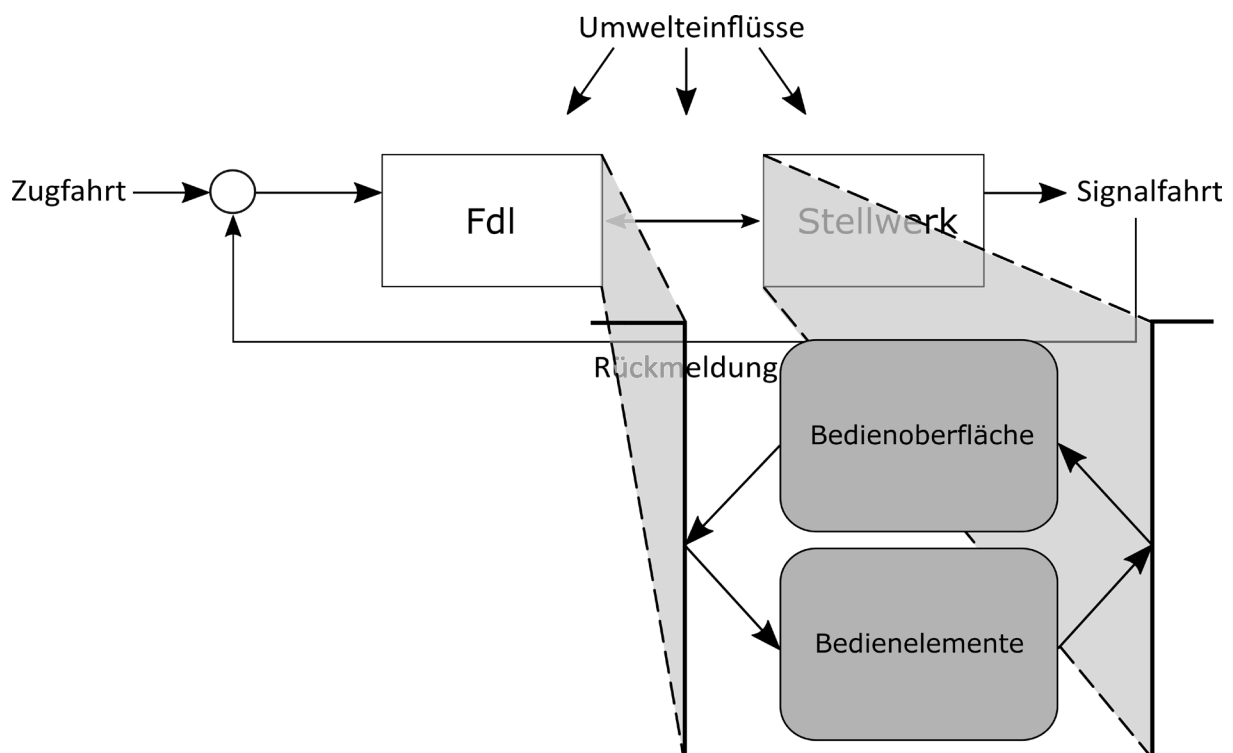


Abbildung 2-9: Lage der Benutzerschnittstelle im modernen Stellwerk (in Anlehnung an [Vol15])

²⁷ Die Benutzerschnittstelle beschreibt die Teile einer Maschine, mit denen Informationen an den Menschen übermittelt werden und über die der Mensch die Maschine bedienen kann [Vol15]. Für die Erforschung der Benutzerschnittstelle von Rechnersystemen hat sich im allgemeinen Forschungsfeld der Ergonomie, die Softwareergonomie als spezielle Wissenschaft entwickelt. Die Softwareergonomie untersucht die Auswirkungen der Benutzerschnittstelle auf den Menschen und die Wirkung von Merkmalen der Benutzerschnittstelle auf die Benutzbarkeit in interaktiven, rechnerbasierten Systemen [Wan93]. Die Ergebnisse sollen dabei helfen, Computersysteme in benutzer- und anwendungsgerechter Weise zu konzipieren, zu realisieren und zu testen [Her09].

2.2.1.3.2 Ingenieurpsychologie

Neben der rein ergonomischen Betrachtungsweise befasst sich auch die angewandte Psychologie mit der MMI. Die Ingenieurpsychologie (IP) untersucht die kognitiven Prozesse, die die Interaktion zwischen dem Menschen und der Maschine prägen. *ZIMOLONG* definiert die IP wie folgt ([Zim06], S. 15):

„Das Ziel der ingenieurpsychologischen Gestaltung ist allgemein die Verbesserung der Funktionalität von Systemen, insbesondere Verbesserungen hinsichtlich der Handhabbarkeit und Gebrauchstauglichkeit, der Zuverlässigkeit, Sicherheit und des Komforts und hinsichtlich der Wirtschaftlichkeit und Umweltverträglichkeit.“

Im Fokus steht also nicht die Funktion des technischen Systems, sondern der Mensch während der Bedienung. Die IP untersucht die menschliche Handlungssteuerung, um die Bedienung so einfach und angenehm wie möglich zu gestalten und ggf. entsprechend zu verbessern. Hierfür untersucht die IP das menschliche Verhalten während der Benutzung des Systems und verwendet dieses Wissen für die Gestaltung der Bedienelemente und der Benutzerschnittstelle.

Exkurs psychologische Modelle

Zur Beschreibung der psychologischen Prozesse dient eine Vielzahl von Modellen. Hierbei sind die Wahrnehmung und das Verständnis von Informationen ein wesentlicher Bestandteil, weil sie eine zentrale Rolle in der Informationsverarbeitung des Menschen einnehmen. Im Folgenden werden zwei Modelle exemplarisch vorgestellt, welche die eben genannten Aspekte betrachten:

Das SEEV-Modell eignet sich zur Beschreibung der Ausrichtung der visuellen Aufmerksamkeit. Zudem schafft es ein Grundverständnis über das Zusammenspiel von Wahrnehmung und Aufmerksamkeit [WHG+01], [WGH+03]. Hierfür integriert das Modell die zwei Stufen der visuellen Wahrnehmung (aufnehmen, selektieren) und macht deutlich, wie äußere Einflüsse und bewusste Strategien die Aufmerksamkeitsverteilung des Menschen steuern [Vol15]. Demnach richtet der Mensch seine Aufmerksamkeit instinktiv auf auffällige Objekte (*Saliency*), wobei ein zu hoher Bewegungsaufwand (*Effort*) vermieden wird. Hinzu kommt, dass der Mensch gezielt Objekte beachtet, von denen er neue Informationen erwartet (*Expectancy*) und die für ihn relevant erscheinen (*Value*).

Das Modell der dreidimensionalen multiplen Ressourcen ist dazu geeignet, die Auswirkung einer Mehrfachaufgabe auf die Leistungsfähigkeit in der Einzelaufgabe zu beschreiben. Es geht davon aus, dass statt einer zentralen Ressource eine Vielzahl von Ressourcen für die Aufnahme und die Verarbeitung der wachsenden Informationsflut vorhanden sind. Diese Ressourcen können gleichzeitig genutzt werden, sind aber jede für sich begrenzt [Wic84]. Greifen Aufgaben auf dieselbe Ressource zu, beeinflusst die Leistung der einen Aufgabe zwangsläufig auch die Leistung der anderen Aufgabe.

2.2.1.3.3 Abgrenzung zwischen Ergonomie und IP

Im Vordergrund beider Wissenschaften steht die schnelle, frustfreie und vor allem fehlerfreie Bedienung der Benutzerschnittstelle [Vol15]. Obwohl sich der Forschungsinhalt beider Wissenschaften in diesem Ziel überschneidet, unterscheiden sie sich in essentiellen Punkten voneinander. Die Ergonomie stützt sich bei der Optimierung der Benutzerschnittstelle auch auf physiologische und anthropologische Eigenschaften des Menschen. Der Schwerpunkt liegt letztendlich auf dem (guten) Produkt. Die Ingenieurpsychologie hingegen befasst sich mit dem besseren Verständnis der kognitiven Prozesse und verbindet den technischen mit dem psychologischen Blickwinkel. Sie ist an der praktischen, lebensnahen Anwendung interessiert [Vol15] und versucht Bedienfehler durch intelligentes Design zu verhindern [BHL08].

2.2.2 Ansätze zur Bewertung der Benutzerfreundlichkeit

Egal wie fortschrittlich die Systeme sind, sie können sich nicht durchsetzen, wenn sie nicht von den Anwendern akzeptiert werden. Deshalb spielt die Benutzerfreundlichkeit bei neuen Techniken eine entscheidende Rolle. Eine hohe Benutzerfreundlichkeit kann die Systemleistung verbessern, die Sicherheit nachhaltig steigern und die Zuverlässigkeit erhöhen. Zudem verringern sich die Risiken, die auf die Gesundheit des Bedieners Einfluss haben [DIN11a]. Um die Benutzerfreundlichkeit eines Systems zu bewerten, gibt es zahlreiche Ansätze. Nachfolgend werden zwei Ansätze vorgestellt, die sich in der Praxis bewährt haben.

2.2.2.1 Usability

Ein wichtiges und umfassendes Konzept für die Bewertung der Benutzerfreundlichkeit liefert die Usability (Gebrauchstauglichkeit). Sie beschreibt leistungsorientierte und aufgabenbezogene Ziele und definiert die Eckpfeiler, die nötig sind, um eine Aussage über die Nutzbarkeit und die Nutzungsqualität treffen zu können. Die Usability eines Produkts wird, vereinfacht gesagt, durch „[...] *the absence of frustration in using it.*“ ([RuCh08], S. 4) erreicht. Im Folgenden wird ein international akzeptierter Standard erläutert, welcher sich in den letzten Jahren etabliert hat, um die Usability zu beschreiben. Dieser Standard ist in einer Deutschen Industrienorm (DIN), der DIN EN ISO 9241, festgeschrieben.

2.2.2.1.1 Grundsatz

In der DIN EN ISO 9241 werden neben einer Definition der Usability, auch Anforderungskataloge, Gestaltungsgrundsätze und Evaluationsmethoden veröffentlicht [DIN99]. Nach der Norm, beschreibt die Usability das Ausmaß, in dem ein System, ein Produkt oder eine Dienstleistung durch bestimmte Benutzer in einem bestimmten Nutzungskontext genutzt werden kann [DIN11a]. Aus dieser Definition ergeben sich drei primäre Aspekte: Effektivität, Effizienz und Zufriedenstellung. Die Effektivität bezieht sich auf die Genauigkeit und die Vollständigkeit der Zielerreichung und der Grad der Zielerreichung bezogen auf den Aufwand wird über die Effizienz beschrieben. Der dritte Aspekt besagt, dass der Benutzer auch subjektiv den Eindruck haben soll, dass sein Handeln effektiv und effizient ist, was unter der Zufriedenstellung zusammengefasst wird.

Weiter schreibt die Norm vor, dass Anzeigen so funktionieren müssen, dass die Wahrscheinlichkeit menschlicher Fehler minimiert wird. Sie sind so auszuwählen, zu gestalten und anzuordnen, dass sie den Merkmalen der menschlichen Wahrnehmung und der durchzuführenden Aufgabe ähnlich sind [DIN14]. Teil 12 der DIN EN ISO 9241 greift diese Forderung auf, fasst charakteristische Eigenschaften dargestellter Information zusammen und gibt Empfehlungen zur deren Darstellung [DIN00].

2.2.2.1.2 Die sieben Grundsätze der Dialoggestaltung

Um die primären Aspekte und die Vorgaben aus der Norm bei der Entwicklung neuer Systeme von Anfang an mit einbeziehen zu können, wird das Konzept der Usability durch sieben Grundsätze der Dialoggestaltung (sieben Dialogprinzipien) unterstützt. Diese Zusammenstellung allgemeiner Ziele dient als Grundlage für die Gestaltung und Bewertung der Usability [DIN08]. Tabelle 2-1 zeigt die sieben Grundsätze und erklärt deren Bedeutung:

Tabelle 2-1: Die sieben Grundsätze der Dialoggestaltung

Gestaltungsprinzip aus DIN EN ISO 9241 Teil 110	Beschreibung (Auszug aus der Norm)
Aufgabenangemessenheit	<ul style="list-style-type: none"> • nur benötigte Informationen sollten angezeigt werden • Eingabe- und Ausgabemedien sollten auf die Belange angepasst sein • Standardwerte sollten vom System vorgegeben werden
Selbstbeschreibungsfähigkeit	<ul style="list-style-type: none"> • angezeigte Informationen sollten den Benutzer anleiten • externe Informationen sollten während der Benutzung minimiert werden • Informationen über erwartete Eingabe und erforderliche Formate sollten bereitgestellt werden
Erwartungskonformität	<ul style="list-style-type: none"> • Interaktionsverhalten und Informationsdarstellung sollten innerhalb eines Systems einheitlich sein • bei Abweichungen von der erwarteten Antwortzeit, sollte der Benutzer hiervon unterrichtet werden • Informationen sollten so strukturiert und organisiert werden, dass sie den Empfindungen des Benutzers entsprechen
Lernförderlichkeit	<ul style="list-style-type: none"> • Rückmeldungen und Erläuterungen sollten den Benutzer unterstützen • das interaktive System sollte es dem Benutzer ermöglichen, die Arbeitsaufgabe mit minimalem Lernaufwand auszuführen
Steuerbarkeit	<ul style="list-style-type: none"> • der Benutzer sollte die Steuerung darüber haben, wie der Dialog abläuft • Beendigung oder Unterbrechung der Arbeit sollte jederzeit ohne Datenverlust möglich sein • die alternative Nutzung von Eingabegeräten sollte, dort wo es geeignet ist, möglich sein
Fehlertoleranz	<ul style="list-style-type: none"> • das interaktive System sollte den Benutzer dabei unterstützen, Eingabefehler zu entdecken und zu vermeiden • die zur Fehlerbehebung erforderlichen Schritte sollten minimiert sein • Fehlermeldungen sollten aussagekräftig sein
Individualisierbarkeit	<ul style="list-style-type: none"> • der Umfang von Erläuterungen sollte dem Wissen des Benutzers entsprechend veränderbar sein • die Informationsdarstellung und die Eingabegeräte sollten an die individuellen Belange des Benutzers angepasst werden können

2.2.2.1.3 Messmethode

Zur Messung der Usability haben sich Fragebögen als (quantitative) Messmethode bewährt. Eine Vielzahl von standardisierten Fragebögen wird Mitte der 1990er und Anfang der 2000er Jahre entwickelt. Der Vollständigkeit halber werden an dieser Stelle die gängigsten deutschsprachigen Usability-Fragebögen genannt: der ISONORM 9241/10 [Prü97], der IsoMetrics [WGH96] und das ErgoNorm-Prüfverfahren [DHF+01].²⁸ Der Vorteil an diesen Fragebögen ist, dass sie die sieben Grundsätze der Gestaltung operationalisiert haben. Es existieren selbstverständlich zahlreiche weitere standardisierte Usability-Fragebögen, auf die hier aber nicht weiter eingegangen wird.

²⁸ Eine Studie mit einer Gegenüberstellung der beiden erst genannten Fragebögen findet sich in [Fig09].

2.2.2.1.4 Kritik

Ein Großteil der Usability-Fragebögen konzentriert sich traditionell auf die Effektivität und die Effizienz der Interaktionen. Beide Aspekte lassen sich gut mit diesem standardisierten Verfahren messen. Anders verhält es sich mit dem dritten Aspekt, der Zufriedenstellung. Die Messung der Zufriedenstellung führt oft zu Diskussionen, da hier Uneinigkeit bzgl. der Wahl der besten Kriterien zur Bewertung besteht [HPB+00]. Es ist unklar, welche Attribute in Betracht gezogen werden müssen, um die Zufriedenstellung ausreichend zu erfassen. Im Teil 11 der DIN EN ISO 9241 heißt es hierzu, dass die Zufriedenstellung durch subjektive Einstufungen auf Skalen oder andere Maßnahmen spezifiziert und gemessen werden kann. An dieser Stelle wird das Verhältnis von positiven zu negativen Äußerungen genannt. Weiter werden Messdaten genannt, welche durch Langzeitmessungen von Merkmalen gewonnen werden können (z. B. Krankheitstage). Das Problem daran ist, dass viele dieser Daten nur in einer direkten Arbeitsumgebung aufgezeichnet werden können, aber nicht in den vielen anderen Situationen, in denen interaktive Systeme verwendet werden. Auch der Datenschutz der Mitarbeiter spielt dabei eine Rolle. Diese Uneinigkeiten führten in den letzten 15 Jahren vermehrt zu der Aussage, dass die Zufriedenstellung im Usability-Ansatz nicht ausreichend betrachtet wird ([HPB+00], S. 202):

“We are aware that user satisfaction is a part of the usability concept provided by ISO 9241-11. However, it seems as if satisfaction is conceived as a consequence of user experienced effectiveness and efficiency rather than a design goal in itself. This implies that ensuring efficiency and effectiveness alone guarantees user satisfaction.”

Demnach stehen im Mittelpunkt der Zufriedenstellungsbetrachtung leistungsorientierte und aufgabenbezogene Ziele. Da es weitere Faktoren gibt, die die wahrgenommene Qualität eines Systems beeinflussen, ist es fragwürdig, ob die Benutzerfreundlichkeit mithilfe des Usability-Ansatzes ausreichend erfasst und bewertet werden kann.

2.2.2.2 User Experience

Effektivität, Effizienz und Zufriedenstellung sind wichtige Faktoren, jedoch gibt es weitere Aspekte, die einen Einfluss auf die wahrgenommene Qualität haben. Der Benutzer verspürt zudem das Bedürfnis nach Stimulation und Identität, d. h. er möchte am Prozess beteiligt werden, herausgefordert werden und sich mit dem Produkt identifizieren [HBK03]. Vorrangig geht es um eine positive Einstellung und die Hoffnung auf Spaß während der Benutzung. Daher muss die klassische Betrachtung weit über die Benutzeranforderungen hinausgehen und auf eine allgemeinere Erlebnisperspektive hinzielen [HaTr06]. Neue Ansätze diskutieren weitere subjektive Bewertungskriterien, wie Ästhetik oder Spaß bei der Nutzung (*joy of use*). Um diese Faktoren mit einbeziehen zu können und die Schwächen der reinen Usability-Betrachtung auszumerzen, wurde der Begriff *User Experience* eingeführt.

2.2.2.2.1 Grundsatz

Die User Experience (UX) weicht von der leistungsorientierten und aufgabenbezogenen Betrachtung ab und beschreibt das Konzept des Nutzungserlebens in seiner Gänze. Dabei finden weiche Bewertungsfaktoren, wie die Ästhetik, im Modell Beachtung [HPB+00]. *NORMAN UND DRAPER* sehen hier die „[...] ultimately central question [...]“ und das wichtigste Kriterium der menschengerechten Gestaltung [NoDr86]. Im Vergleich zur reinen Zufriedenstellung nach DIN EN ISO 9241 beschreibt die UX nicht nur ein messbares Ergebnis, sondern einen komplexen Prozess. *HASSENZAHL* beschreibt die UX mithilfe von drei Qualitäten. Dazu gehören die Attraktivität, die pragmatische und die hedonische Qualität [HPB+00]. Dabei beschreibt die Attraktivität die Akzeptanz bzw. die Ablehnung der Benutzer gegenüber dem Produkt global. Bei der pragmatischen Qualität handelt es sich im Kern um die Aspekte der klassischen Usability. Hierzu gehören u. a. Steuerbarkeit, Effektivität, Effizienz und Erlernbarkeit. Stimulation, Ästhetik und Identifikation mit dem Produkt gehören zur hedonischen Qualität. Abbildung 2-10 zeigt die drei Qualitäten. Zu beachten ist, dass zwischen den drei Qualitäten keine hierarchische Ordnung besteht. Die Attraktivität bezieht sich auf das gesamte Produkt und nicht auf einzelne Aspekte, weshalb sie als übergeordnete Qualität dargestellt wird. Unabhängig davon werden alle Qualitäten als gleichwertig gesehen.

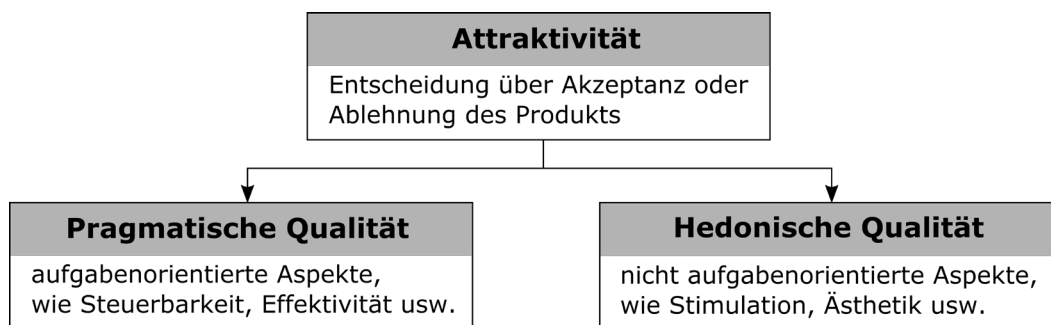


Abbildung 2-10: Zusammenhang der drei UX-Qualitäten nach *HASSENZAHL* (in Anlehnung an [RSP+13])

Besonders die Beachtung der hedonischen Qualität ist nicht zu vernachlässigen. Beispielsweise hat der wahrgenommene Spaß während der Nutzung einen positiven Effekt auf die Akzeptanz [Mos12] und die Zufriedenstellung der Benutzer [ISW94]. Hinzu kommt, dass die wahrgenommene Ästhetik einen Einfluss auf die subjektive Beurteilung der Usability hat [TKI00]²⁹, was die Leistungsbereitschaft der Benutzer erhöhen kann [MoTh10]. Diese positiven Auswirkungen sprechen definitiv für eine stärker Beachtung der hedonischen Qualität.

2.2.2.2.2 Strukturbild

Im *Components of User Experience Model* (CUE-Modell) werden die drei UX-Qualitäten integriert und der logische Ablauf der Bewertung eines interaktiven Systems veranschaulicht (Abbildung 2-11). Das interaktive System setzt sich aus den technischen Spezifikationen, relevanten Eigenschaften des Benutzers und der zu erledigenden Aufgabe zusammen [THN09].

²⁹ siehe hierzu auch [ISH08] und [ISH09]

Die Wahrnehmung der unterschiedlichen Qualitäten führt direkt zur Bewertung des Systems und indirekt zu emotionalen Reaktionen. Die Emotionen beeinflussen ebenfalls die Nutzungsqualität und somit auch die Systembewertung.³⁰ Im rechten Teil der Abbildung wurde, durch den Verfasser der vorliegenden Arbeit, ein Pfad für die Leistungsbereitschaft eingefügt. Der Verfasser möchte damit zum Ausdruck bringen, dass sich eine positive Bewertung des interaktiven Systems durch den Benutzer, in einer höheren Akzeptanz dem System gegenüber niederschlägt. Diese höhere Akzeptanz führt wiederum zu einer höheren Leistungsbereitschaft des Benutzers.

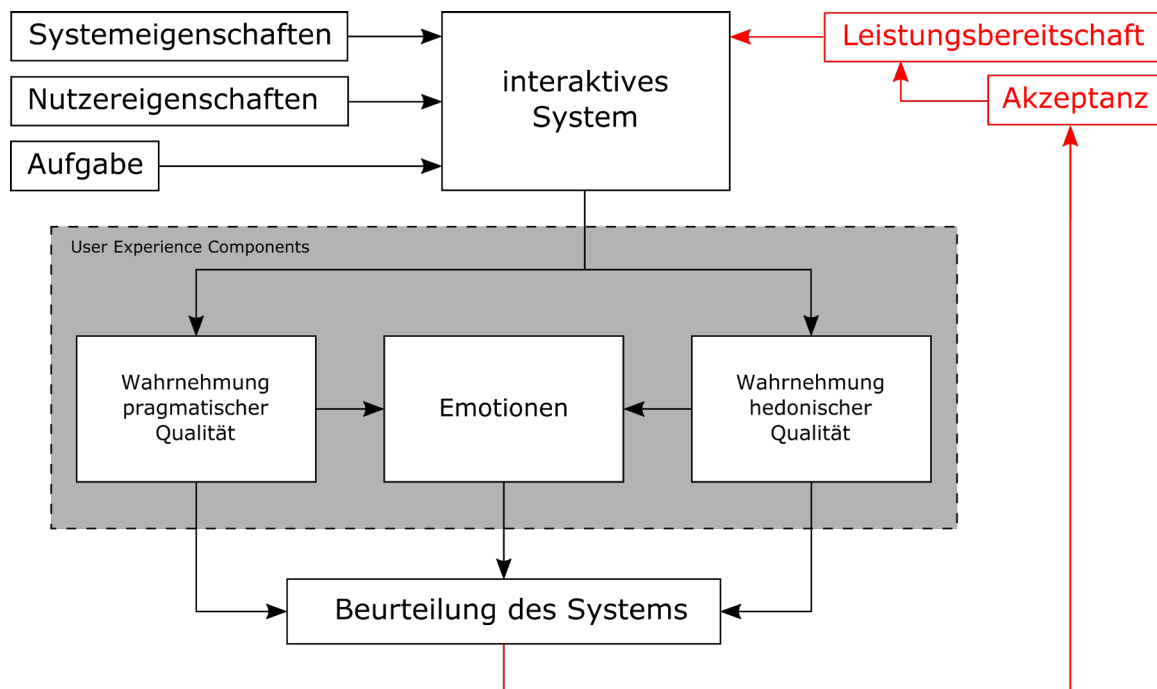


Abbildung 2-11: CUE-Modell (in Anlehnung an [ThMa07], vereinfacht und erweitert durch den Verfasser)

2.2.2.2.3 Messmethode

Da die pragmatische Qualität der Usability entspricht, kann sie standardisiert gemessen werden. Für die hedonische Qualität gab es solche Standardverfahren anfangs noch nicht, was ein möglicher Grund für die Vernachlässigung dieser Faktoren bei der Bewertung von Bedienoberflächen ist [Lau01]. In den letzten Jahren wurden eine Reihe von Methoden zur Evaluation und Bewertung der hedonischen Qualität entwickelt. Wie auch bei dem Qualitätsmerkmal *Usability*, eignen sich besonders Fragebögen für eine schnelle und unkomplizierte Erhebung der UX.³¹ Die gängigsten Fragebögen, welche sich bereits in der Praxis bewährt haben, sind der AttrakDiff2 [HBK03] und der *User Experience Questionnaire* (UEQ) [LSH06], [LHS08].³² Beide Verfahren benutzen Bewertungsskalen (Rating-Skalen), um die Qualität unterschiedlicher Testobjekte ermitteln zu können (Kapitel 5.1.3.2). Im Vergleich zu klassischen Usability-Fragebögen haben die UX-Fragenbögen den Vorteil, dass sie neben der

³⁰ Für weiterführende Informationen wird auf [Mah06] verwiesen.

³¹ Einen umfassenden Überblick über weitere Methoden gibt [Mah08].

³² Als Beispiel für den AttrakDiff2 sei auf [SHL04] und für den UEQ auf [LSI+09] sowie [IHS09] verwiesen.

pragmatischen auch die hedonische Qualität abfragen und bewerten. Hierdurch werden alle Wahrnehmungen und Reaktionen, die aus der tatsächlichen und/oder der erwarteten Benutzung eines Produkts, eines Systems oder einer Dienstleistung resultieren, untersucht.

2.2.3 Zusammenfassung Kapitel 2.2

Der Arbeitsplatz des Fdl hat sich dahingehend gewandelt, dass die Arbeit im modernen Stellwerk im Wesentlichen am Bildschirm stattfindet. Der Vorteil daran ist, dass hierdurch Informationen flexibel dargestellt, räumlich konzentriert und verschiedene Systeme integriert werden können. Zugleich birgt dies jedoch die Gefahr, die Informationsaufnahme-fähigkeit des Bedieners zu überfordern, da die auf dem Bildschirm dargestellten Informationen eine sehr hohe Dichte haben. Gerade in Gefahrensituationen oder bei Störungen sind die rasche Entdeckung und die Identifikation der Störquelle zur Entscheidungsfindung notwendig [Wid90]. Aus diesem Grund wird der Interaktion zwischen Mensch und Maschine besondere Beachtung geschenkt. Für die Untersuchung der Interaktion zwischen beiden Akteuren ist es wichtig den Informationsaustausch zu verstehen. Hierfür muss die Mensch-Maschine-Interaktion analysiert werden.

Im Wesentlichen befassen sich zwei Wissenschaften mit der Erforschung und Verbesserung der Mensch-Maschine-Interaktion. Es handelt sich dabei um die Ergonomie, als Teilgebiet der Arbeitswissenschaft, und die Ingenieurpsychologie, als Forschungsbereich der angewandten Psychologie. Im Fokus beider Wissenschaften liegt die Benutzerschnittstelle. Die Benutzerschnittstelle beschreibt die Teile einer Maschine, mit denen Informationen an den Menschen übermittelt werden und über die der Mensch die Maschine bedienen kann. Im modernen Stellwerk ist die Benutzerschnittstelle der Bedienplatz. Der Unterschied zwischen beiden Wissenschaften ist das Ziel der Analysen. Die Ergonomie stützt sich bei der Optimierung der Benutzerschnittstelle auch auf physiologische und anthropologische Eigenschaften des Menschen und legt den Schwerpunkt eher auf das fertige Produkt. Die Ingenieurpsychologie hingegen befasst sich mit dem besseren Verständnis der kognitiven Prozesse und ist an der praktischen, lebensnahen Anwendung interessiert. Durch die Optimierung der Benutzerschnittstelle soll die Benutzerfreundlichkeit erhöht werden, welche als Maß für die erlebte Nutzungsqualität gesehen wird.

Zur Bewertung der Benutzerfreundlichkeit eignen sich zwei Ansätze: die Usability und die User Experience. Die Usability wird in der DIN EN ISO 9241 beschrieben. Die Usability eines Produkts setzt sich aus den Aspekten *Effektivität*, *Effizienz* und *Zufriedenstellung* zusammen. In die Kritik geraten ist der Ansatz, weil der Fokus auf einer sehr leistungsorientierten und aufgabenbezogenen Betrachtung liegt. Deshalb ist der Ansatz der User Experience besser geeignet, um eine umfassendere Bewertung des Gesamtsystems durchzuführen. Die User Experience erweitert den klassischen Usability-Ansatz um weitere Faktoren, wie *Attraktivität*, *Stimulation*, *Ästhetik* und *Identifikation*. Zur Messung beider Ansätze haben sich Fragebögen als praktikable Möglichkeit etabliert.

2.3 Einfluss von Farbe und Kontrast

Die User Experience bewertet neben der pragmatischen auch die hedonische Qualität. Letztere steht in der vorliegenden Arbeit besonders im Fokus, da Untersuchungen gezeigt haben, dass diese weichen Faktoren den Umgang mit der Bedienoberfläche positiv beeinflussen können. Beispielsweise kann die wahrgenommene Ästhetik, die subjektive Wahrnehmung der Usability erhöhen [TKI00]. Dadurch wird die Akzeptanz des Bedieners gegenüber dem System erhöht [Mos12] und der emotionale Zustand des Bedieners wird stabilisiert. Das wiederum fördert die Flexibilität beim Problemlösen und regt das kreative Denken an, wodurch die Leistungsfähigkeit des Bedieners zusätzlich noch verbessert werden kann.

Ziel dieses Kapitels ist die Vorstellung zweier Attribute, welche einen wesentlichen Einfluss auf die Bewertung der Ästhetik und somit auf die subjektive Wahrnehmung der Usability des Bedienplatzes haben. Des Weiteren wird im zweiten Teil dieses Kapitels ein wissenschaftlicher Ansatz vorgestellt, mit dem die Bedienoberfläche eines Stellwerks in einer ersten Näherung auf ihre Benutzerfreundlichkeit überprüft werden kann.

2.3.1 Vorstellung der Attribute Farbe und Kontrast

Ästhetische Reize, wie Farbe und Kontrast, sind wichtige Attribute für die Bewertung der Bedienoberfläche. Dabei beschreibt die Farbe einen aus der Umwelt aufgenommenen Farbreiz, der im Gehirn als Farbe empfunden wird und der Kontrast das Helligkeitsverhältnis zweier Farben [BiSc10]. Im weiteren Verlauf werden weitere Eigenschaften der Farbe und des Kontrasts beschrieben.

2.3.1.1 Farbe

Eine Farbe wird durch die Eigenschaften Farbton, Helligkeit und Sättigung beschrieben [BiSc10]. Sie kommt zum Einsatz, um Informationen zu kodieren, zu gliedern und zu gewichten. Zudem kann sie verwendet werden, um Aufmerksamkeit zu erregen und um die Fehlererkennung zu vereinfachen [DIN12]. Dabei ist noch zu beachten, dass, wenn Informationen rein über eine farbliche Kodierung übermittelt werden, diese nicht mehr als fünf Ausprägungen haben sollten [Mil56]. Das bedeutet, wenn beispielsweise die Ausleuchtung eines Signalsymbols auf der Bedienoberfläche eine Information über mehrere Zustände beinhaltet, hierbei nicht mehr als fünf verschiedene Farben verwendet werden sollten.

Der gezielte Einsatz von Farbe kann sich positiv auf die visuelle Informationsverarbeitung auswirken [Chr75], [DaOs88] und spielt bei der Objektwahrnehmung eine bedeutsame Rolle [WLL+04].³³ Hierbei ist jedoch zu beachten, dass die Anzahl der gleichzeitig auf einer Anzeige

³³ Die Objektwahrnehmung kann zusätzlich durch eine Blinkkodierung erhöht werden [Chr75]. Diese sollte aber nur außerordentlich wichtigen Informationen vorbehalten bleiben, um die alarmierende Wirkung nicht aufzuheben. Die Kombination von Blink- und Farbkodierung kann diesen Effekt nochmals verstärken [Col90].

dargestellten Farben, auf den Leistungsanforderungen der Arbeitsaufgabe und den Leistungsvoraussetzungen der Bediener beruhen. Deshalb sollten auf dem Bildschirm nur so viele Informationen wie gerade nötig dargestellt werden. Hieraus ergibt sich, dass der Standardfarbensatz (einschließlich weiß, grau und schwarz) aus nicht mehr als elf Farben besteht [CaRe78]. Die DIN EN ISO 9241 geht noch ein Stück weiter und legt fest, dass, wenn schnelle visuelle Suchvorgänge erforderlich sind, nicht mehr als sechs Farben verwendet werden sollten [DIN12].

2.3.1.2 Kontrast

Die Kontrastwirkung verschiedener Farben ist ein weiterer wichtiger Punkt, der einen Einfluss auf die Bildschirmarbeit hat. Die Fähigkeit Kontraste wahrzunehmen ist notwendig, um Objekte zu erkennen. Durch den Kontrast wird die Sichtbarkeit von Informationen gewährleistet [WLL+04]. Das heißt, je höher der Kontrast, desto besser sind Objekte voneinander zu unterscheiden. Das hängt mit der Sehschärfe zusammen. Je höher die Sehschärfe ist, desto leichter können z. B. Gefahrenquellen erkannt werden.

Um die Augen durch intensive optische Reize nicht zu stark zu belasten und von der eigentlichen Aufgabe abzulenken, sollten helle, wenig gesättigte Farben für Hintergründe verwendet werden. Dabei bieten achromatische (unbunte) Kontraste die beste Lesbarkeit bei gleichzeitiger Schonung der Augen. Besonders empfehlenswert sind Kombinationen aus leicht grauen Hintergründen mit schwarzer oder sehr dunkler Schriftfarbe. Der Vorteil an dieser Kombination ist, dass die Hintergründe fast weiß erscheinen, die dunkleren Farben allerdings nicht so stark überstrahlt werden und der korrekte Farbeindruck dadurch erhalten bleibt [Hol07].

Des Weiteren kann die Wirkung des Kontrasts auf dem Bildschirm genutzt werden, um benachbarte Farbflächen voneinander abzugrenzen und die ästhetische Bewertung zu verbessern [Lau01]. Weiße oder schwarze Linien bzw. Flächen können die Kontrastwirkung auf dem Bildschirm noch verstärken. Dabei lässt Weiß alle Farben eher dunkler und weniger leuchtend wirken, während Schwarz die Leuchtkraft steigert und die Farbe heller erscheinen lässt [Hol07]. Das ist auch der Grund, warum dunkle Schrift auf hellem Hintergrund schmaler und zierlicher erscheint als umgekehrt.³⁴

³⁴ Bei Druckerzeugnissen verhält es sich im Übrigen umgekehrt [Hol07].

2.3.2 Vorgaben für barrierefreie Informationstechnik

Bei der Gestaltung einer Bedienoberfläche muss die Farbe sehr gezielt und bewusst eingesetzt werden. Dieser Ansatz ist wichtig, wird aber oft vernachlässigt. Aus diesem Grund wird in der vorliegenden Arbeit ein Konzept gesucht, um die Bedienoberflächen vorab zu bewerten und eine Aussage über deren Benutzerfreundlichkeit zu treffen.

Eine Möglichkeit, die farbliche Gestaltung³⁵ einer Bedienoberfläche zu bewerten, wird in dem Konzept der barrierefreien Informationstechnik gesehen. Der Gesetzgeber schreibt vor, dass Informationstechnik, ob am Arbeitsplatz oder unterwegs, barrierefrei und für alle Menschen gleichermaßen zugänglich sein muss [BUND07]. Diese Forderung gilt am Fahrdienstleiterarbeitsplatz nur bedingt, da konzernrechtliche Bestimmungen u. a. die Farbenblindheit im Betriebsdienst nicht zulassen. Nichtsdestotrotz gibt es eine Reihe von Vorschlägen, die bei einer Neuentwicklung einer Bedienoberfläche mit ins Kalkül einbezogen werden sollten, um die Benutzbarkeit zu verbessern. Das *World Wide Web Consortium* (W3C) ist Herausgeber der *Web Content Accessibility Guidelines* (WCAG)³⁶. In den Richtlinien werden Vorgaben für barrierefreie Webinhalte gemacht, mit dem Ziel, die Zugänglichkeit von interaktiven Systemen für alle Personengruppen zu erhöhen [W3C09]. Die Vorgaben gelten nicht nur für Menschen mit Behinderung, sondern tragen allgemein zur besseren Erkennbarkeit bei. Aus diesem Grund hat der Gesetzgeber die Forderungen der WCAG 2.0 in eine eigene Verordnung überführt [BUND11].

2.3.2.1 Der RGB-Wert

Um den Kontrast berechnen zu können, muss die relative Farbhelligkeit der einzelnen Farben bestimmt werden. Hierfür wird der RGB-Wert benötigt. Dieser spiegelt die Koordinate des Vektors der Farbvalenz³⁷ im RGB-Farbraum wider und zeigt den Anteil der Grundfarben in der gemischten Farbe. Der RGB-Farbraum dient als Grundlage zur Darstellung von Farbbildern auf Bildschirmen. Er setzt sich aus den drei Grundfarben **R**ot, **G**rün und **B**lau zusammen und ermöglicht die additive Mischung (fast) aller Farben.

Die Primärfarbvalenzen repräsentieren die Einheitsvektoren mit den Koordinaten für Rot (1, 0, 0), Grün (0, 1, 0) und Blau (0, 0, 1). Der Wert 1 entspricht dem höchsten Grundfarbenanteil und der Wert 0 steht für den niedrigsten Anteil. Die Farbe *Schwarz* besitzt die Koordinaten (0, 0, 0) und die Farbe *Weiß* die Koordinaten (1, 1, 1). Das heißt, dass sich die Farbe *Weiß* aus 100 % der drei Grundfarben zusammensetzt und deshalb die hellste aller Farben im RGB-Farbraum ist. Um auf die in der Rechnertechnik üblichen Werte von null bis 255 zu kommen, werden jeder Primärfarbvalenz acht Bit zugeordnet. Dadurch kann jede Grundfarbe 256 Sättigungswerte annehmen.

³⁵ beinhaltet die Farbe und den Kontrast

³⁶ Aktuell befinden sich die WCAG in der Version 2.0.

³⁷ Eine Farbe im Farbraum wird als Farbvalenz bezeichnet und kann im RGB-Farbraum über den RGB-Wert genau beschrieben werden.

2.3.2.2 Bestimmung der relativen Farbhelligkeit

Die relative Farbhelligkeit (FH) beschreibt wie hell eine Farbe auf dem Bildschirm erscheinen wird. Mit der Formel 2-1 aus Abbildung 2-12 wird die relative Farbhelligkeit einer Farbe im RGB-Farbraum berechnet [W3C08]. Dabei liefert die angegebene Formel 2-1 nur eine Prognose darüber, weil die Helligkeit auch von der verwendeten Technik und den Einstellungen des Bildschirms abhängig ist.³⁸

$$FH = 0.2126 * \left(\frac{R}{255} \right)^{2.2} + 0.7152 * \left(\frac{G}{255} \right)^{2.2} + 0.0722 * \left(\frac{B}{255} \right)^{2.2}$$

Abbildung 2-12: Bestimmung der relativen Farbhelligkeit (Formel 2-1)

Die relative Farbhelligkeit kann einen Wert zwischen null und eins annehmen. Hierbei entspricht der Wert 0 der Farbe *Schwarz* und der Wert 1 der Farbe *Weiß*. Hieraus ergibt sich, dass je heller die Farbe ist, desto näher kommt die Farbhelligkeit dem Wert 1. Die Umrechnung der relativen Farbhelligkeit in den RGB-Farbwert erfolgt durch die Multiplikation der relativen Farbhelligkeit mit dem maximalen RGB-Endwert von 255.

2.3.2.3 Bestimmung des Kontrasts

Der Kontrast beschreibt das Verhältnis zwischen den Helligkeiten zweier Farben. Er wird mit der Formel 2-2 aus Abbildung 2-13 errechnet [W3C08]:

$$\text{Kontrast} = \frac{(L1 + 0.05)}{(L2 + 0.05)}$$

Abbildung 2-13: Bestimmung des Kontrasts (Formel 2-2)

L1 steht für die relative Farbhelligkeit der helleren Farbe und L2 für die relative Farbhelligkeit der dunkleren Farbe. Wie bereits erwähnt, können die Helligkeitswerte einen Wert von null (für die Farbe *Schwarz*) bis maximal eins (für die Farbe *Weiß*) erreichen. Zu beachten ist, dass der Kontrast immer als Verhältnis angegeben wird. Abbildung 2-14 zeigt den errechneten Kontrast für die Farben *Schwarz* und *Weiß*. Dieser beträgt 21, was für ein Verhältnis von 21:1 steht. Der Einfachheit halber wird an dieser Stelle der Begriff *Kontrastverhältnis* eingeführt und im weiteren Verlauf der Arbeit verwendet.³⁹

$$\text{Kontrast}_{\text{schwarz/weiß}} = \frac{(L_{\text{weiß}} + 0.05)}{(L_{\text{schwarz}} + 0.05)} = \frac{(1 + 0.05)}{(0 + 0.05)} = 21$$

Abbildung 2-14: Kontrast der Farben *Schwarz* und *Weiß*

³⁸ Auf eine umfassende Erklärung der Formel wird an dieser Stelle verzichtet und auf [NN10] verwiesen.

³⁹ Diese Begrifflichkeit entspricht nicht ganz der Semantik, genaugenommen beschreibt das Kontrastverhältnis das Verhältnis zweier Kontraste. Diese Trennung ist für das Verständnis der Untersuchung nicht notwendig.

2.3.2.4 Weitere Vorgaben des W3C

Das W3C legt weitere Grenzwerte fest. Hierzu zählen die Konformitätsstufe, die Helligkeitsdifferenz zwischen Vorder- und Hintergrundfarbe und der Farbunterschied zwischen beiden Farben.

2.3.2.4.1 Konformitätsstufen

Das errechnete Kontrastverhältnis (KV) wird benötigt, um das Softwareprodukt in die Konformitätsstufen (KS) nach [W3C09] einzuteilen. Tabelle 2-2 zeigt die vier Stufen:

Tabelle 2-2: Stufen der Konformität der WCAG 2.0

Stufe	Schriftgrad	Kontrastverhältnis (KV)
AA	größer 18 pt oder 14 pt (Fett)	3:1
AA+	regulär	4.5:1
AAA	größer 18 pt oder 14 pt (Fett)	4.5:1
AAA+	regulär	7:1

Die Stufen ergeben sich aus dem verwendeten Schriftgrad und dem Kontrastverhältnis zwischen Vorder- und Hintergrundfarbe. Mindestanforderung für barrierefreie Webinhalte ist die Stufe AA. Diese wird erreicht, wenn der verwendete Schriftgrad mindestens 18 pt oder 14 pt (Fett) beträgt und das Kontrastverhältnis einen Wert von 3:1 erreicht (Abbildung 2-15).

Diese Darstellung entspricht der
Konformitätsstufe AA

Abbildung 2-15: Beispiel für Konformitätsstufe AA, mit Schriftgrad 14 pt (Fett) und einem Kontrastverhältnis 3:1 [Hintergrund-RGB-Wert (149, 149, 149), Vordergrund-RGB-Wert (255, 255, 255)]

Um die Konformitätsstufe AA+ zu erreichen, muss das Kontrastverhältnis der Darstellung einen Wert von 4.5:1 haben. Dabei spielt der Schriftgrad nur indirekt eine Rolle (Abbildung 2-16).

Diese Darstellung entspricht der
Konformitätsstufe AA+

Abbildung 2-16: Beispiel für Konformitätsstufe AA+, mit Schriftgrad 14 pt und einem Kontrastverhältnis 4.5:1 [Hintergrund-RGB-Wert (118, 118, 118), Vordergrund-RGB-Wert (255, 255, 255)]

Wird in der Abbildung 2-16 der Schriftgrad auf mindestens 18 pt oder 14 pt (Fett) erhöht, erreicht das Beispiel die Stufe AAA.

2.3.2.4.2 Helligkeitsdifferenz

Die Helligkeitsdifferenz (HD) beschreibt die Differenz zwischen der Helligkeit des Hintergrunds und der Helligkeit des Vordergrunds. Sie kann einen Wert zwischen null und 255 annehmen. Dabei muss, um eine ausreichende Sichtbarkeit beider Farben zu gewährleisten, die Differenz größer als 125 sein [W3C00]. Der Wert wird mit der Formel 2-3 aus Abbildung 2-17 errechnet:

$$\begin{aligned} \text{HD} &= \text{helle Farbe} - \text{dunkle Farbe} \\ &= \left[\frac{(R_{\text{Hell}} * 299) + (G_{\text{Hell}} * 587) + (B_{\text{Hell}} * 114)}{1000} \right] \\ &\quad - \left[\frac{(R_{\text{Dunkel}} * 299) + (G_{\text{Dunkel}} * 587) + (B_{\text{Dunkel}} * 114)}{1000} \right] \end{aligned}$$

Abbildung 2-17: Bestimmung der Helligkeitsdifferenz (Formel 2-3)

2.3.2.4.3 Farbunterschied

Eine weitere Möglichkeit, um eine gute Sichtbarkeit zu gewährleisten, ist ein ausreichend großer Farbunterschied (FU) zwischen den verwendeten Farben. Der Farbunterschied kann einen Wert zwischen null und 765 annehmen und sollte einen Mindestwert von 500 haben [W3C00]. Er wird mit Formel 2-4 aus Abbildung 2-18 errechnet:

$$\begin{aligned} \text{FU} &= [\max(R1 \text{ o. } R2) - \min(R1 \text{ o. } R2)] + [\max(G1 \text{ o. } G2) - \min(G1 \text{ o. } G2)] \\ &\quad + [\max(B1 \text{ o. } B2) - \min(B1 \text{ o. } B2)] \end{aligned}$$

Abbildung 2-18: Bestimmung des Farbunterschieds (Formel 2-4)

Die Formel 2-4 bezieht sich ausschließlich auf die RGB-Werte. Unabhängig von den relativen Farbhelligkeiten beider Farben, wird immer der kleinere Wert vom größeren abgezogen.

2.3.2.5 Tools zur Überprüfung der Vorgaben des W3C

Um die Vorgaben der WCAG 2.0 zu überprüfen, gibt es eine Reihe von webbasierten Programmen. Diese eignen sich sehr gut, um Vorder- und Hintergrundfarbe miteinander zu vergleichen.⁴⁰ Die Tools führen die Rechnungen und Kontrollen automatisiert durch und geben die Ergebnisse für Helligkeitsdifferenz, Farbunterschied, Kontrast und Konformitätsstufe aus.⁴¹

⁴⁰ In der vorliegenden Arbeit wurde der *Colour Contrast Check* verwendet [NN05]. Dieses Tool bietet vielfältige Einstellmöglichkeiten, um die bestmögliche Farbkombination zu finden.

⁴¹ Die webbasierten Programme arbeiten in der Regel mit der hexadezimalen Darstellung des RGB-Werts. Für die Umrechnung des RGB-Werts in die hexadezimale Darstellung wurde [NN0] verwendet.

2.3.3 Zusammenfassung Kapitel 2.3

Eine Möglichkeit die subjektive Wahrnehmung der Usability positiv zu beeinflussen, ist die Verwendung einer ästhetisch wirkenden Bedienoberfläche. Diese Anforderung kann durch ein auf die Situation abgestimmtes Farbkonzept erfüllt werden. Zwei wichtige Aspekte sind hierbei die Farbe und der Kontrast. Eine Farbe ergibt sich aus einem aus der Umwelt aufgenommenen Farbreiz und der Kontrast beschreibt das Verhältnis zwischen zwei Farbhelligkeiten. Beide Aspekte haben diverse Eigenschaften, welche von Vorteil sein können. So eignet sich die Farbe beispielsweise dazu Aufmerksamkeit zu lenken und mithilfe des Kontrasts kann die Objektwahrnehmung unterstützt werden.

Auf der Suche nach einem Konzept für eine erste Bewertung der Anwendbarkeit der Bedienoberfläche wurde in dem Konzept der barrierefreien Informationstechnik ein Näherungsverfahren gefunden. Das Konzept macht Vorgaben für barrierefreie Webinhalte mit dem Ziel, die Zugänglichkeit von interaktiven Systemen zu verbessern und die Erkennbarkeit von Objekten zu erhöhen. Diese Vorgaben werden von dem *World Wide Web Consortium* in den *Web Content Accessibility Guidelines* beschrieben. Neben der Erläuterung der Vorgabe, werden in diesen Richtlinien auch Messgrößen definiert, um diese Vorgaben zu überprüfen. Hierzu gehören die relative Farbhelligkeit, der Kontrast, die Konformitätsstufe, die Helligkeitsdifferenz und der Farbunterschied. Grundlage für alle Messgrößen ist der RGB-Wert der Farbe.

2.4 Die Darstellungsarten

Das grundsätzliche farbliche Konzept eines Bildschirms wird als Darstellungsart bezeichnet. Unabhängig vom Darstellungsverfahren (CRT, LCD, LED usw.) wird zwischen der Positiv- und der Negativdarstellung unterschieden. Bei der Darstellung von dunklen Zeichen auf hellem Untergrund wird, egal ob ein- oder mehrfarbig, von der Positivdarstellung gesprochen. Die Darstellung von hellen Zeichen auf dunklem Untergrund wird als Negativdarstellung bezeichnet.

In der Ursprungsform wird bei der Positivdarstellung ein weißer Hintergrund mit schwarzer Schrift verwendet und bei der Negativdarstellung weiße Schrift auf schwarzem Hintergrund. Abbildung 2-19 zeigt beide Darstellungsarten. Für eine bessere Unterscheidbarkeit wird das Beispiel ausschließlich in achromatischen Farben dargestellt.

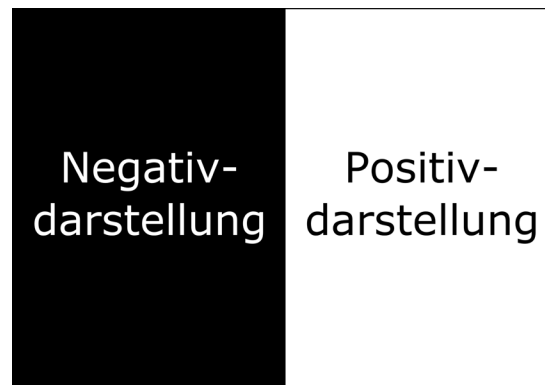


Abbildung 2-19: Positiv- vs. Negativdarstellung mit achromatischen Farben

2.4.1 Gegenüberstellung

Die Wahl der Darstellungsart sollte in Abhängigkeit von den Beleuchtungsverhältnissen der Umgebung erfolgen [BAuA12], wobei eine Erhöhung der Beleuchtungsstärke im Arbeitsbereich zu einer Leistungssteigerung führen kann [Com87]. Das ist ganz klar ein Vorteil, welcher für die Positivdarstellung spricht. Durch den helleren Hintergrund ermöglicht die Positivdarstellung eine höhere Beleuchtungsstärke der Umgebungsbeleuchtung. Das gewährleistet eine Anwendung in Räumen mit natürlicher Umgebungsbeleuchtung [BAuA12]. Geringe Ermüdung, gute Arbeitsqualität und ein niedriges Unfallrisiko sind mögliche Folgen. Die Darstellungsart entspricht im Kern der Darstellung von Zeichen auf Papiervorlagen. Dadurch mindert sie im Vergleich zu einem dunklen Hintergrund die Gefahr störender Lichtreflexionen und Blendungen. Zudem bietet sie bessere Anpassungsmöglichkeiten der Augen an die Arbeitsumgebung [Com87]. Es werden unnötige Hell-Dunkel-Adaptionsvorgänge vermieden [CaRe78]. Demnach ist mit der Positivdarstellung eine deutlich günstigere Voraussetzung für die Lesbarkeit der Anzeigen gegeben [BAuA14].

Die Negativdarstellung hingegen erlaubt wegen des höheren Risikos von Spiegelungen, Reflexionen oder Streulichteinflüssen keine höhere Beleuchtungsstärke der Umgebungsbeleuchtung [BAuA14]. Sie sollte nur in abgedunkelten Räumen verwendet werden. Dies erschwert das Lesen von nicht-bildschirmgebundenen Medien wie z. B. Papier. Der ständige Hintergrundwechsel zwischen heller Vorlage und dunklem Bildschirm sowie die damit verbundenen Adaptionsvorgänge im Auge begünstigen die Ermüdung des Betrachters, was Unbehagen und eine Leistungsverringerung hervorrufen kann [Com87]. Auf der anderen Seite eignet sich die Negativdarstellung besonders dann, wenn Farben zur Kodierung von Informationen eingesetzt werden. Sie besitzt in diesem speziellen Fall eine bessere Kontrastwirkung. In Tabelle 2-3 werden die Vor- und Nachteile beider Darstellungsarten zusammengefasst und gegenübergestellt.

Tabelle 2-3: Gegenüberstellung der Vor- und Nachteile der betrachteten Darstellungsarten

	Positivdarstellung	Negativdarstellung
Vorteil	<ul style="list-style-type: none"> • kann in Räumen mit Tageslicht verwendet werden, höhere Umgebungsbeleuchtung führt zur Leistungssteigerung • verringert störende Lichtreflexionen und Blendungen • bessere Lesbarkeit bei kleinen Displays und kleiner Schrift • altersunabhängige bessere Korrekturleseleistung • bessere Anpassungsmöglichkeit der Augen, verringert Ermüdungserscheinungen 	<ul style="list-style-type: none"> • bessere Kontrastwirkung von Hintergrund und Zeichen, begünstigt die Kodierung von Informationen und die Fehlererkennung
Nachteil	<ul style="list-style-type: none"> • Gefahr des Flimmerns • Gefahr, dass der Hintergrund die anderen Farben überstrahlt, ungenügender Kontrast 	<ul style="list-style-type: none"> • Gefahr von störenden Reflexionen und Blendungen durch die Umgebungsbeleuchtung • sollte nur in abgedunkelten Räumen verwendet werden • vermehrt Hell-Dunkel-Adaption, begünstigt Ermüdungserscheinungen

2.4.2 Einführung der Polarität

Moderne Bedienoberflächen verwenden nicht nur achromatische (unbunte) Farben. Aus diesem Grund ist die Einteilung in Positiv- und Negativdarstellung nicht immer direkt möglich. Ein Beispiel hierfür ist der Originalbedienplatz der in der Studie verwendeten Stadtbahnsimulation.⁴² Augenscheinlich verwendet der Bedienplatz einen helleren Hintergrund, was die Vermutung zulässt, dass es sich um eine Positivdarstellung der Bedienoberfläche handelt. Der Großteil der Vordergrundfarben besitzt jedoch eine höhere relative Farbhelligkeit als die Hintergrundfarbe, was eher der Negativdarstellung entspricht (Kapitel 4.2.4). Um die Darstellungsart für die in der Studie verwendeten Bedienoberflächen zu ermitteln, wurde nach einer passenderen Einteilung gesucht.

Einen Ansatz liefert das Konzept aus dem Teil 303 der DIN EN ISO 9241. Die Norm weicht die klassische Aufteilung auf und unterscheidet ausschließlich nach dem farblichen Konzept der Bedienoberfläche. Sie unterteilt die Bedienoberflächen in zwei Gruppen. Zum einen in Bedienoberflächen mit achromatischem Hintergrund (unbunt) hinter chromatischen Vordergrundbildfarben (bunt) und zum anderen in Bedienoberflächen mit achromatischen Vordergrundbildfarben auf einem chromatischen Hintergrund [DIN12]. Dadurch werden die Begriffe *Positiv-* und *Negativdarstellung* obsolet. Ein Problem dieses Konzepts ist, dass die relative Farbhelligkeit keine Beachtung findet. Demnach spielt es keine Rolle, welche Helligkeiten die Vorder- bzw. die Hintergrundfarben haben. Sie könnten bei den einzelnen Informationen

⁴² siehe hierzu Kapitel 4.2.1, Abbildung 4-1

sehr verschieden sein, wodurch zwar die Definition angewendet werden kann, aber dem Anspruch an ein ergonomisches Bedienoberflächendesign nicht Genüge getan ist. Deshalb ist es, aus Sicht des Verfassers der vorliegenden Arbeit, notwendig, die relativen Farbhelligkeiten bei der Unterscheidung einzubeziehen.

In der englischsprachigen Literatur werden für die unterschiedlichen Darstellungsarten die Begriffe *positive (display) polarity* und *negative (display) polarity* verwendet. Durch den Begriff *Polarität* lässt sich die klassische Einteilung der Positiv- und der Negativdarstellung mit der Unterscheidung aus [DIN12] vereinen. Hat das Farbkonzept eine positive Polarität bedeutet das, dass die Vordergrundfarben dunkler sind als die Hintergrundfarben. Umkehrt sind bei der negativen Polarität die Hintergrundfarben dunkler als die Vordergrundfarben. Dabei spielt es keine Rolle, ob bunte Farben auf unbunten Hintergründen dargestellt werden oder unbunte Farben auf bunten Hintergründen. Aus diesem Grund wird dieser Ansatz in der vorliegenden Arbeit forciert und in der Bedienoberflächenstudie verwendet, um die Darstellungsart der verschiedenen Bedienplätze der Stadtbahnsimulation zu bestimmen (Kapitel 4.2.4).

2.4.3 Aktuelle Situation bei der DB Netz AG

Stelltische und Panoramatafeln der Relaisstellwerke verwenden eine positive Polarität. Im Gegensatz dazu, sind im ESTW der DB Netz AG die Bedienoberflächen mit negativer Polarität vorherrschend. Dieser Zustand ändert sich auch mit der Einführung des iBS nicht. Das erste ESTW ging im Jahre 1985 in Betrieb. Demnach ist die Bedientechnik über 30 Jahre alt. Ein Grund für die Bevorzugung der negativen Polarität in den 80er Jahren ist, dass die damaligen Röhrenmonitore nicht in der Lage waren, visuelle Abläufe flimmerfrei mit einer positiven Polarität wiederzugeben [CaRe78]. In heutigen Betriebsleitzentralen dominieren LCD-Monitore das Umfeld und die technischen Voraussetzungen für den Einsatz von Anzeigen mit positiver Polarität sind gegeben.

Bis dato erfolgten keine großen Veränderungen im Design der ESTW-Bedienoberflächen. Die Gründe hierfür sind vielfältig. So liegen der Technik lange Produktlebenszyklen zu Grunde und proprietäre Lösungen erschweren die Umsetzung moderner Anforderungen [Kon13]. Hinzu kommt, dass die Erneuerung der Alttechnik aus wirtschaftlichen Gesichtspunkten nur sukzessive geschieht und das Arbeitskonzept heutiger Betriebszentralen eher gedimmte Beleuchtungsverhältnisse vorsieht, wodurch bessere Voraussetzungen für Bildschirme mit negativer Polarität geschaffen sind. Zudem werden Innovationen dadurch erschwert, dass eine Mischung verschiedener Bedienoberflächen verhindert werden soll und die Anhebung der Beleuchtungsstärke sowie die Veränderung des Beleuchtungskonzepts mit erheblichem Aufwand verbunden sind. Ein weiterer Grund ist, dass der Bediener die Möglichkeit haben soll mit einer bestimmten Farbe einen bestimmten Sachverhalt zu verbinden. Das heißt, die Farben bzw. Farbkombinationen, die sich bereits bei existierenden Bedienoberflächen durchgesetzt haben, sollen beibehalten werden.

2.4.4 Zusammenfassung Kapitel 2.4

Die Darstellungsart beschreibt das grundsätzliche farbliche Konzept eines Bildschirms. Dabei wird zwischen der Positiv- und der Negativdarstellung unterschieden. Der Vergleich beider Darstellungsarten zeigt, dass die Nachteile der Negativdarstellung überwiegen und aus ergonomischer Sicht die Positivdarstellung bei modernen Bedienoberflächen bevorzugt verwendet werden sollte. Werden jedoch Farben zur Kodierung von Informationen eingesetzt, ist die Negativdarstellung besser geeignet. In diesem Fall ist die Kontrastwirkung zwischen Vorder- und Hintergrundfarbe besser als bei der Positivdarstellung.

Moderne Bedienoberflächen verwenden nicht nur achromatische Farben. Dadurch ist die Einteilung in Positiv- und Negativdarstellung nicht immer direkt möglich – was der Originalbedienplatz der Stadtbahnsimulation zeigt. Aus diesem Grund wurde im Rahmen der Arbeit nach einer neuen Einteilung gesucht. Abhilfe wird durch die Einführung des Begriffs *Polarität* geschaffen. Bei der Bestimmung der Polarität werden die relativen Farbhelligkeiten der Vorder- und Hintergrundfarben miteinander verglichen.

Historisch bedingt verwendet die DB Netz AG in allen ESTW-Generationen mit Bildschirm die negative Polarität als bevorzugte Darstellungsart. Eine Änderung dieses Zustands ist auch in neuen Stellwerksgenerationen nicht zu erkennen, wie die Entwicklung des iBS zeigt. Ob dieses Vorgehen noch zeitgemäß ist und welche Wirkung die farbliche Darstellung auf den Bediener hat, wird in der vorliegenden Arbeit untersucht.

3 ZIELSETZUNG DER ARBEIT

Aus ergonomischer Sicht haben Bedienoberflächen mit negativer Polarität bedeutend mehr Nachteile als Bedienoberflächen mit positiver Polarität (Tabelle 2-3). Das ist ein Grund, warum einige Infrastrukturbetreiber in ihren neuen Stellwerksgenerationen eine Bedienoberfläche mit positiver Polarität verwenden. Dieses Vorgehen kann durch Publikationen aus dem englischsprachigen Raum zusätzlich bekräftigt werden. Hier werden in Studien die kognitiven Auswirkungen der Darstellungsart auf den Bediener untersucht. In diesen Untersuchungen schneiden Bedienoberflächen mit positiver Polarität besser ab [PMM+13]. Sie kommen zu dem Ergebnis, dass die positive Polarität, besonders bei kleineren Displays oder textbasierter Kommunikation mit kleiner Schrift, zu empfehlen ist [PMB14]. Die Ergebnisse scheinen für sich zu sprechen. Demnach wären Bedienoberflächen mit positiver Polarität zu favorisieren.

Bei genauerer Betrachtung der zahlreichen Untersuchungen fällt ins Auge, dass die Ergebnisse recht heterogen sind und kein eindeutiges Bild liefern. Gerade bei Themen wie Fehlererkennung und Lesbarkeit von Zeichen, weichen die gewonnenen Erkenntnisse stark voneinander ab. So gibt es Studien, die für die Verwendung der positiven Polarität sprechen (exemplarisch [SDJ+90]) und wiederum andere, die für die Verwendung der negativen Polarität sprechen (exemplarisch [CPW78], [VAC05]).

Um die Ursachen für solche Abweichungen besser verstehen zu können, werden in der vorliegenden Arbeit nicht die verschiedenen Darstellungsarten der Bedienoberflächen miteinander verglichen – vielmehr wird ein rudimentärerer Ansatz verfolgt. Die zentrale Fragestellung der Arbeit befasst sich mit der Untersuchung des Einflusses der farblichen Gestaltung auf den Bediener. Die Literatur zeigt, dass es aus theoretischen Gründen sinnvoll ist, hier einen Effekt zu erwarten (Kapitel 2.3). Aus diesem Grund werden in der Studie zwei Bedienoberflächen miteinander verglichen. Die untersuchten Bedienoberflächen unterscheiden sich nur hinsichtlich der farblichen Gestaltung, nicht aber in Bezug auf Benutzbarkeit und Funktionalität, voneinander.

Nachfolgend werden die Ziele und Fragestellungen der Arbeit im Detail vorgestellt. Der Fokus liegt hierbei auf der Erklärung des Ursprungs der Überlegungen. Eine umfassende Beschreibung der abhängigen Variablen und deren Operationalisierung, wird in Kapitel 5 behandelt.

3.1 Einfluss der farblichen Gestaltung

Der erste Teil der Bedienoberflächenstudie befasst sich mit der Untersuchung des Einflusses der farblichen Gestaltung der Bedienoberfläche auf den Stellwerksbediener. Hierfür werden verschiedene Variablen erfasst, welche in den folgenden Kapiteln vorgestellt werden.

3.1.1 User Experience

Ausgangspunkt für die Idee, die UX bei der Bewertung der Bedienoberflächen ins Kalkül einzubeziehen, ist die Aussage, dass die UX als eine Erweiterung der klassischen Usability gesehen werden kann (Kapitel 2.2.2.2). Weiche Faktoren, wie sie im UX-Konzept betrachtet werden, können einen nicht unerheblichen Einfluss auf den Benutzer des interaktiven Systems haben. Im Vergleich zu den Methoden der Usability, die bereits angewendet wurden, um die Usability einer Stellwerksbedienoberfläche zu bewerten [MLB15], gibt es bis dato keine öffentlichen empirischen Untersuchungen, die den UX-Ansatz auf das Eisenbahnwesen anwenden.

Als Grundlage für die Untersuchung der User Experience dient das Konzept von *HASSENZAHL* [HPB+00]. Dieses Konzept befasst sich mit der Attraktivität sowie mit der pragmatischen und der hedonischen Qualität eines Produkts (Kapitel 2.2.2.2). Im Rahmen der Bedienoberflächenstudie werden die drei Qualitäten mithilfe des User Experience Questionnaire (UEQ) erfasst und anschließend die Ergebnisse der untersuchten Bedienoberflächen miteinander verglichen. Dabei werden zwei Ansätze verfolgt. Zum einen wird überprüft, ob sich der UEQ für die Messung der UX einer Bedienoberfläche im Stellwerk eignet und zum anderen, ob mit diesem Verfahren Änderungen am farblichen Konzept erfasst und nachgewiesen werden können. Des Weiteren ist von Interesse, ob eine hohe pragmatische Qualität die Leistungsfähigkeit und die Handlungssicherheit der Versuchsteilnehmer positiv beeinflusst.

3.1.2 Psychische Beanspruchung

Wie in Kapitel 2.2 beschrieben, wird eine mangelhafte Anpassung der farblichen Gestaltung an menschliche Fähigkeiten als störend und belastend empfunden. Dies kann zu einer psychischen Beanspruchung führen, was sich indirekt auf die Arbeitsbereitschaft auswirkt und wodurch die Sicherheit des Gesamtsystems beeinflusst wird. Aus diesem Grund schließt die Gestaltung eines sicheren Systems die Berücksichtigung psychischer Beanspruchung ein [NeTi76]. Im Rahmen der Studie wird die psychische Beanspruchung der Versuchsteilnehmer während der Eingewöhnungsphase aufgezeichnet (Kapitel 5.1.5.2).

Ziel ist es, die geistige und die körperliche Beanspruchung während des Erlernens der Simulation zu erfassen, mögliche Unterschiede zwischen den verwendeten Bedienoberflächen

aufzuzeigen und den Einfluss der psychischen Beanspruchung auf andere Variablen im Experiment zu untersuchen.

3.1.3 Befindlichkeit

In der Bedienoberflächenstudie wird untersucht, ob die farbliche Gestaltung der Bedienoberfläche einen Einfluss auf die Stimmung bzw. die Befindlichkeit⁴³ hat. Dabei kennzeichnet die Befindlichkeit das nicht beobachtbare Verhalten eines Individuums [SSN+94]. Nach [Lau01] kann die Befindlichkeit durch die drei Dimensionen *Gesundheit*, *Wachheit* und *Entspannung* gesamthaft beschrieben werden. Dies entspricht weitestgehend dem *Strukturmodell der Stimmungen* nach [Sch99]. Ursprung der Überlegung, die Stimmung während des Versuchs zu erfassen, ist, dass die Stimmung des Bedieners vor allem dann von Bedeutung ist, wenn durch sie gesundheitliche Probleme entstehen. Beispielsweise ist eine länger anhaltende, niedergedrückte Stimmung ein wesentlicher Bestandteil von Depressionen, eine durch starke Unruhe gekennzeichnete Befindlichkeit geht mit Angststörungen einher und eine andauernde Müdigkeit ist charakteristisch für das chronische Erschöpfungssyndrom [HDP+12]. Deshalb ist die Betrachtung der Stimmung auch für das Arbeiten im Stellwerk wichtig.

In der vorliegenden Arbeit wird die Befindlichkeit der Versuchsteilnehmer erfasst, um mögliche Ursachen für fehlerhaftes Verhalten aufzuzeigen. Dabei spielt vor allem die farbliche Gestaltung der Bedienoberfläche eine Rolle, weil nachgewiesen wurde, dass sie einen positiven Einfluss auf die Aktivierung der Benutzer haben kann [PfWi86], [PfWi87]. Dieser Effekt soll auch im Bedienoberflächenkontext nachgewiesen werden. Zudem wird untersucht, ob die anderen Dimensionen von der verwendeten Bedienoberfläche beeinflusst werden und welchen Einfluss die Befindlichkeit auf andere Variablen im Experiment hat.

3.1.4 Bedienungssicherheit

In anderen Industriezweigen wurde herausgefunden, dass eine farblich-ästhetische Gestaltung der Bedienoberfläche die subjektive Wahrnehmung der eigenen Bedienungssicherheit erhöhen kann [Wan93]. Im Rahmen der Bedienoberflächenstudie wird diese Feststellung auf das Eisenbahnwesen übertragen und auf ihre Validität überprüft.

Des Weiteren wird vermutet, dass die Bedienungssicherheit einen Einfluss auf das Überwachungsverhalten hat. Die Annahme ist, dass ein Bediener, welcher sich sicher fühlt, zu jedem Zeitpunkt ein besseres Situationsbewusstsein besitzt als einer, der sich unsicher fühlt. Dem trainierten Bediener fällt es leichter, die Betriebssituation trotz Ablenkung schnell und richtig einzuschätzen. Demnach wird angenommen, dass ein gut trainierter Versuchsteil-

⁴³ Strenggenommen beschreibt der Begriff *Befindlichkeit* ein mehrdimensionales Modell und der Begriff *Stimmung* nur eine Dimension dieses Modells. Da es für beide Begriffe keine klare Unterscheidung gibt [HDP+12], werden sie synonym verwendet.

nehmer mehr Reservekapazität⁴⁴ zur Verfügung hat und deshalb der Nebenaufgabe eine höhere Aufmerksamkeit widmet, ohne dabei einen Leistungseinbruch in der Hauptaufgabe zu riskieren und/oder mehr Bedienfehler zu machen.

3.1.5 Fehlerentdeckungszeit

Die Fehlerentdeckungszeit gehört eigentlich zu den vordefinierten Leistungskennzahlen (Kapitel 3.2.1). Der Großteil dieser Kennzahlen wird für den Versuch verwendet, eine Aussage über die Leistungsfähigkeit und die Handlungssicherheit der Versuchsteilnehmer zu treffen. Ziel ist es, die Bedienereffizienz zu operationalisieren. Da bei der Fehlerentdeckungszeit ein Einfluss der farblichen Gestaltung vermutet wird, wird diese Variable bereits an dieser Stelle kurz vorgestellt.

Im Rahmen der Studie wird untersucht, ob sich die farbliche Gestaltung der Bedienoberfläche auf die Fehlerentdeckungszeit der Versuchsteilnehmer auswirkt. Besonders von Interesse sind hierbei Erkenntnisse bezogen auf einen möglichen Einfluss der psychischen Beanspruchung, der Befindlichkeit und der Bedienungssicherheit.

3.2 Operationalisierung der Bedienereffizienz

Ein weiteres Ziele der Arbeit ist es, zu einem besseren Verständnis der Bedienereffizienz im Stellwerksbedienoberflächenkontext beizutragen. Die Ergebnisse sollen dazu dienen, mögliche Ansatzpunkte für die Entwicklung von Methoden zur Messung der Bedienereffizienz abzuleiten. Ausgangspunkt der Bedienereffizienzüberlegungen ist, dass eine effiziente Bedienung zu einer höheren Zufriedenheit des Nutzers führt [FJP11], [JeMe12].

In anderen Industriezweigen ist der Begriff der Bedienereffizienz immer häufiger anzutreffen [Pro11]. Im Eisenbahnwesen verhält es sich anders. Hier sind empirische Forschungsergebnisse eher selten. Das mag daran liegen, dass es für die Bedienereffizienz bislang weder eine allgemein eingeführte Definition, noch verbindliche Messgrößen und Verfahren zur Bewertung gibt. *PACHL* beschreibt die Bedienereffizienz als Fähigkeit, Bedienhandlungen räumlich und zeitlich so zu bündeln, dass sich ein zeitsparender Bedienungsablauf ergibt [Pac15]. Da der Bediener kaum Einfluss auf den Durchsatz in seinem Bedienbereich hat, eignen sich nach [Pac15] durchsatzbezogene Bezugsgrößen im Eisenbahnwesen nur bedingt, um die Bedienereffizienz zu messen.

Diese mögliche Einschränkung wird in der Bedienoberflächenstudie bewusst ignoriert, um eine Basis für die Entwicklung von Methoden zur Messung der Bedienereffizienz im Stellwerk zu schaffen. Da die verwendete Simulation einen festen Fahrplan besitzt, bietet die Studie

⁴⁴ siehe hierzu Kapitel 5.1.3.10

die Möglichkeit, den Durchsatz über alle Versuchsteilnehmer annähernd konstant⁴⁵ zu halten. Dieser Umstand wird genutzt, um vordefinierte Leistungskennzahlen zu erfassen. Diese werden anschließend in einen Zusammenhang gebracht, um die Bedienereffizienz versuchsweise zu operationalisieren.

3.2.1 Übersicht der vordefinierten Leistungskennzahlen

Tabelle 3-1 zeigt eine Übersicht der gesammelten Leistungskennzahlen. Diese Kennzahlen spannen ein Modell auf, welches die Faktoren *Leistungsfähigkeit* und *Handlungssicherheit* beschreiben. Im Folgenden werden die Leistungskennzahlen kurz skizziert. Eine ausführliche Beschreibung des theoretischen Ansatzes und der einzelnen Leistungskennzahlen erfolgt in Kapitel 5.1.4.

Tabelle 3-1: Übersicht der Leistungskennzahlen

Nr.	Aussage	Bezeichnung	Kurzbeschreibung
1	Leistungsfähigkeit	Anzahl der Zugfahrten	Gesamtdurchsatz an Zugfahrten während des Hauptversuchs
2	Leistungsfähigkeit	Fehlerentdeckungszeit	Zeit bis zum Erkennen einer Unregelmäßigkeit
3	Leistungssicherheit	Klick-Fehler-Verhältnis (KFV)	Verhältnis zwischen Klickzahl und gemachten Fehlern in der Nebenaufgabe
4	Handlungssicherheit	Anzahl der Hilfshandlungen	Gesamtanzahl der während des Hauptversuchs getätigten Hilfshandlungen (ohne 3)
5	Handlungssicherheit	Anzahl der Bedienfehler	Gesamtanzahl an unnötigen und falschen Bedienhandlungen
6	Handlungssicherheit	Gesamtüberwachungszeit	Gesamtblickverweildauer der Versuchsteilnehmer auf dem Bildschirm

3.2.2 Untersuchung der Mausbewegung

Wie in [Pac15] gezeigt wird, eignet sich das Mousetracking-Verfahren als Methode zur Bewertung der Bedienereffizienz im Stellwerk. In ersten Fallstudien konnte gezeigt werden, dass mit einem einfachen Tracker, Aussagen über den Trainingsstand und den Einfluss von Komfortfunktionen während der Bedienung getroffen werden können. Im Rahmen der vorliegenden Arbeit wird dieses Verfahren verwendet, um diese Ergebnisse an einer größeren Versuchsteilnehmerzahl zu validieren. Des Weiteren sollen die gesammelten qualitativen Daten mit quantitativen Messwerten in einen Zusammenhang gebracht werden, um die Bedienereffizienz zu operationalisieren. Es wird angenommen, dass sich eine hohe subjektive Einschätzung der eigenen Leistung mit strukturierten Mausbewegungen und größeren Verweilkreisen in Verbindung setzen lässt. Dies soll im Rahmen der durchgeführten Studie empirisch nachgewiesen werden.

⁴⁵ siehe hierzu Kapitel 5.1.4

4 DIE VERSUCHSUMGEBUNG

Die Studie wird im Virtuellen Eisenbahnbetriebslabor (V-EBL) des Instituts für Eisenbahnwesen und Verkehrssicherung (IfEV) durchgeführt. Im V-EBL sind Stellwerkssimulationen unterschiedlicher Betreiber und verschiedener Bauformen vorhanden. Dabei wird der Stelltisch als Bildschirmgrafik dargestellt und die Bedienung erfolgt per Maus.

Ziel dieses Kapitels ist es, die verwendeten Bedienplätze vorzustellen. Es wird die Ausleuchtung und die Bedienung erläutert. Ferner wird die Darstellungsart der verwendeten Bedienplätze definiert und auf die verschiedenen Unregelmäßigkeiten, welche während des Versuchs eintreten können, näher eingegangen.

4.1 Allgemeine Beschreibung des Bedienplatzsystems

Als Versuchsumgebung dient ein Schulungssystem der Firma Siemens. Bei dem System handelt es sich um ein Leitsystem aus der Produktfamilie VICOS® OC (*Vehicle and Infrastructure Control and Operating System/Operation Management System Commuter*). Es besteht aus vier Schülerbedienplätzen⁴⁶ und einem Lehrerbedienplatz. Auf dem Lehrerbedienplatz übernimmt die Trainersoftware VICOS T&S (*VICOS Training and Simulation*) die Simulation der Stellwerke, der Außenanlage und der Züge. Zudem werden auch die Unregelmäßigkeiten über eine Eingabemaske an dem jeweiligen Feldelement eingeschaltet und zurückgenommen.

Das System bildet originale Fahrdienstleiterarbeitsplätze des Typs Vicos OC 100 nach und verfügt mittels Bedienplatz OC 111 über eine verfahrensgesicherte Anzeige⁴⁷. Den Versuchsteilnehmern wird somit ein Arbeitsplatz zur Verfügung gestellt, der in der Bedienung und Ausleuchtung identisch zur Realität ist. Durch diese Konfiguration ist der Versuch praxisnah und zum Testen der verschiedenen Ansätze sehr effektiv.

⁴⁶ Aufgrund methodischer, organisatorischer und technischer Faktoren werden während des Versuchs nur zwei von vier Schülerarbeitsplätzen gleichzeitig verwendet.

⁴⁷ Durch die verfahrensgesicherte Anzeige wird sichergestellt, dass dem Bediener, mit der notwendigen Sicherheit, die tatsächlichen Zustände der Außenanlagen angezeigt werden. Grund hierfür ist, dass der Bediener die Anzeige betrieblich nur auswerten darf, wenn sie signaltechnisch sicher ist [Jon12].

Simuliert wird das Netz einer deutschen Stadtbahn mit integrierter Zuglenkung. Die Zuglenkung ermöglicht einen komplett automatisierten Betrieb, wodurch die Belastung des Bedieners gesenkt und die Leistungsfähigkeit des Systems erhöht wird [BWS+12]. Der Zuglenkbetrieb (ZLB) kann durch den Bediener⁴⁸ jederzeit ausgeschaltet werden. Die Hauptaufgabe der Versuchsteilnehmer während des Experiments besteht in der Überwachung des Bedienbereichs und in dem Erkennen von Unregelmäßigkeiten (Kapitel 4.3.2).

Neben der Überwachung des Bedienbereichs müssen die Versuchsteilnehmer eine Nebenaufgabe erledigen (Kapitel 5.1.3.10). Die Nebenaufgabe besteht aus einer visuellen Suchaufgabe, bei der die Versuchsteilnehmer einen Zielreiz (Target) unter mehreren Ablenkern (Distraktoren) finden müssen. Hierdurch wird eine Mehrfachaufgabenbelastung simuliert und eine visuelle und kognitive Ablenkung erzeugt. Es wird vermutet, dass es durch die Nebenaufgabe zu einem Leistungseinbruch in der Überwachungsaufgabe einiger Versuchsteilnehmer kommt (Kapitel 3.1.4).

4.2 Die verwendeten Bedienplätze der Stadtbahn

Da das komplette Netz der Stadtbahn verhältnismäßig groß ist, ist die Darstellung auf einem Bildschirm nicht möglich. Aus diesem Grund wird das Netz in mehrere Bildschirmausschnitte unterteilt. Diese Ausschnitte werden als Lupenbilder bezeichnet. In der Praxis hat der Bediener mehrere Bildschirme an seinem Arbeitsplatz. An diesen kann der Bediener sich die relevanten Lupenbilder aufschalten. Um die Versuchsteilnehmer in der Studie nicht zu überfordern, wird nur ein Bildschirm mit einem vordefinierten Lupenbild verwendet (exemplarisch Abbildung 4-1). Bei der Auswahl des Bildes wurde Wert auf eine hohe Komplexität gelegt. Auf dem verwendeten Lupenbild muss der Bediener acht Signale, fünf Weichen und eine Kreuzung überwachen. Im Vergleich zu den anderen Lupenbildern, weist sie eine überdurchschnittlich hohe Anzahl verschiedenartiger Feldelemente auf. Hinzu kommt, dass den Versuchsteilnehmern durch das Arbeiten mit nur einem Lupenbild die Möglichkeit gegeben wird, während der Eingewöhnungsphase die Örtlichkeit kennenzulernen und eine ausreichende Bedienungssicherheit für den Hauptversuch zu erlangen. Zudem werden für alle Versuchsteilnehmer gleiche Bedingungen geschaffen, weil sie sich auf ein Lupenbild konzentrieren können und dieselben Voraussetzungen haben.

Nachfolgend werden die verschiedenen Bedienplätze mit den verwendeten Bedienoberflächen vorgestellt. Es wird die Ausleuchtung erklärt und die Polarität der Bedienoberfläche bestimmt. Die dabei verwendeten Abbildungen dienen nur der Unterscheidung und der Unterstützung der Erläuterungen. Eine detailliertere Darstellung der Bedienoberflächen findet sich im Anhang dieser Arbeit (Anhang F).

⁴⁸ Im Folgenden wird ausschließlich die Bezeichnung *Bediener* verwendet, da es sich bei den Versuchsteilnehmern um nicht ausgebildete Fahrdienstleiter handelt.

4.2.1 Original- und Experimentalbedienplatz

Der Originalbedienplatz wird in der Realität bei einer deutschen Stadtbahn verwendet (Abbildung 4-1). Im V-EBL wird er für die universitäre Lehre eingesetzt.

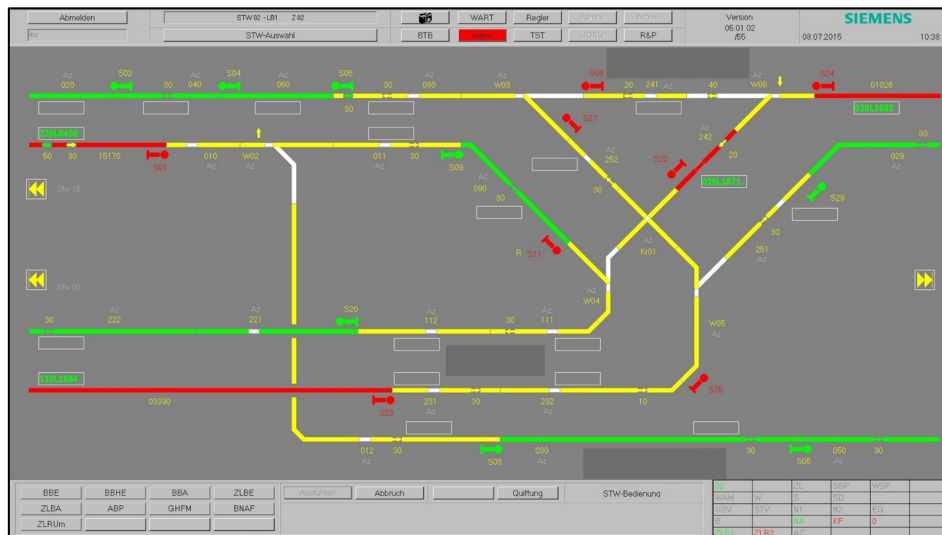


Abbildung 4-1: Bedienoberfläche OC 111 (Originalbedienplatz)

Der zweite Bedienplatz, im Folgenden als Experimentalbedienplatz bezeichnet, wurde durch die Firma Siemens nach den Vorgaben des Verfassers der vorliegenden Arbeit entwickelt (Abbildung 4-2). Ursprung der Entwicklung war eine Voruntersuchung im Rahmen einer Lehrveranstaltung. Das Farbkonzept des Experimentalbedienplatzes orientiert sich an dem einer ESTW-Bedienoberfläche, wie sie aktuell bei der DB Netz AG Anwendung findet. Dabei wurde bewusst ein dunkleres Grün für die Darstellung der eingestellten Fahrstraße verwendet. Das Originalgrün der ESTW-Bedienoberfläche ist sehr viel heller (ähnlich Abbildung 4-1), wodurch der Farbunterschied zwischen Grün und Gelb geringer wäre. Dies würde die Unterscheidbarkeit erschweren und wurde deshalb geändert.

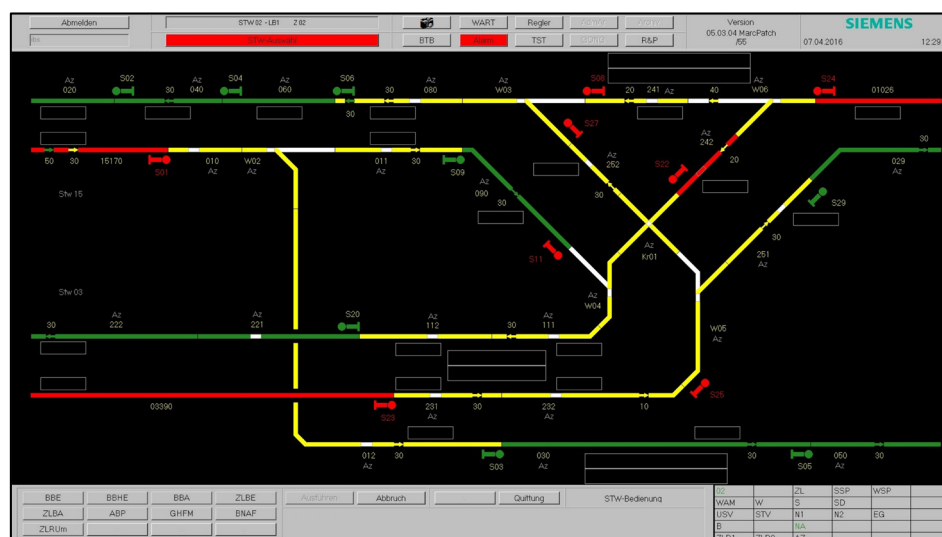


Abbildung 4-2: Bedienoberfläche OC 111 (Experimentalbedienplatz)

4.2.2 Designbedienplatz

Abbildung 4-3 zeigt den Designbedienplatz. Im Gegensatz zu dem Original- und dem Experimentalbedienplatz verwendet der Designbedienplatz eine achromatische (unbunte) Ausleuchtung des Fahrwegs in Grundstellung. Hierdurch wird das Hauptaugenmerk auf die Zustände *Fahrstraße*, *Zugfahrt* und *Störung* gelegt. Des Weiteren ergibt sich ein harmonischeres Bild, was die Orientierung und die Fehlererkennung erleichtern soll.

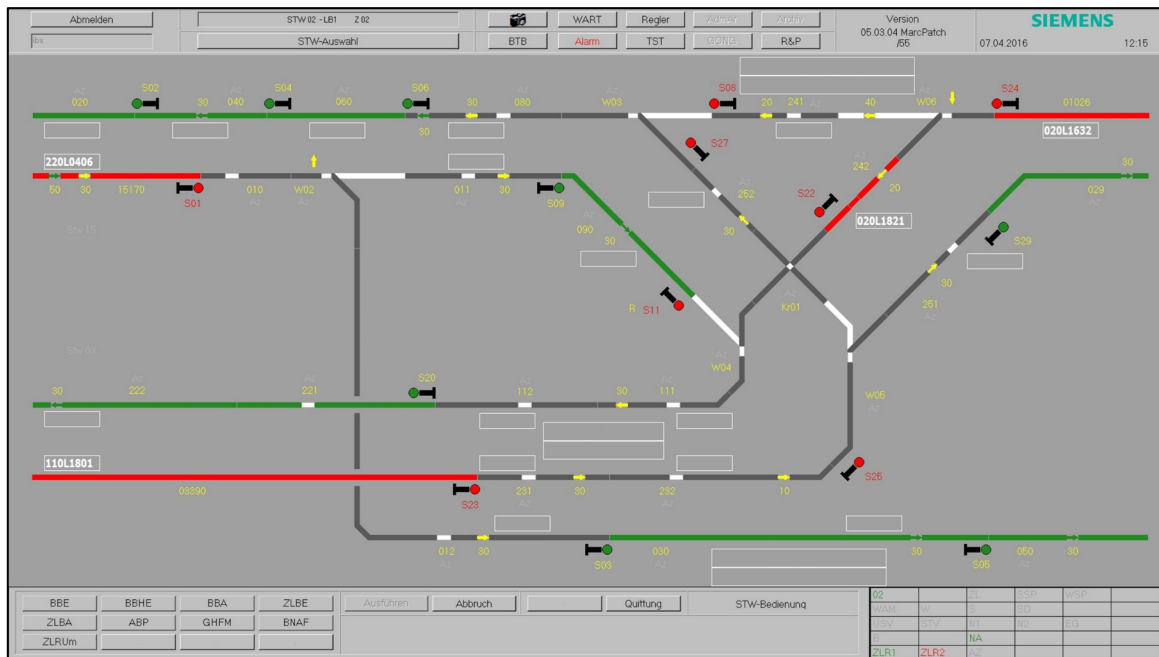


Abbildung 4-3: Bedienoberfläche OC 111 (Designbedienplatz)

Ein weiterer Unterschied zu den anderen Bedienplätzen ist, dass die Signale schwarz sind und nur der Signalkopf eine Ausleuchtung besitzt. Dadurch wird die Kontrastwirkung erhöht und die Signale grenzen sich besser von dem Hintergrund ab (Kapitel 2.3.1.2). Das Problem dabei ist, dass die zusätzliche Information über den Status des Signals augenscheinlich wegfällt, da diese normalerweise über die Ausleuchtung des Signalfußes ausgegeben wird (Kapitel 4.2.3). Um diesen Informationsverlust zu verhindern, hat der Versuchsteilnehmer bei dem Designbedienplatz die Möglichkeit, sich den Status des Signals über ein Pop-up-Fenster ausgegeben zu lassen. Das Fenster wird durch Anklicken des jeweiligen Signalfußes eingeblendet. Abbildung 4-4 zeigt exemplarisch einige Meldungen.

Trotz der Änderungen wird das grundsätzlich bestehende Farbkonzept weiterhin genutzt. Fahrstraßen werden immer noch grün ausgeleuchtet und Zugfahrten und Störungen der Gleisfreimeldung rot. Der Designbedienplatz wurde ebenfalls für die Voruntersuchung entwickelt. Aufgrund der zu Beginn der Bedienoberflächenstudie eingeschränkten Funktionalität des Designbedienplatzes konnte dieser nicht für den Hauptversuch verwendet werden. Er wird von den Versuchsteilnehmern im Rahmen des Experiments nur qualitativ und nicht quantitativ bewertet.

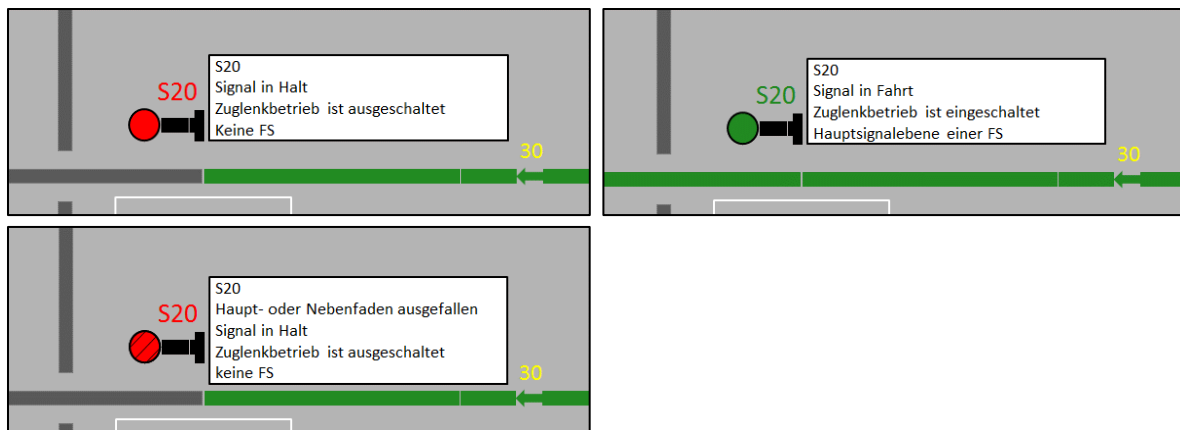


Abbildung 4-4: Darstellung weiterer Informationen über ein Pop-up-Fenster (oben links: Signal in Halt, ZLB aus, keine Fahrstraße (FS)/oben rechts: Signal in Fahrt, ZLB ein, FS eingestellt/unten links: Signal in Halt, ZLB aus, keine FS, Signalstörung)

4.2.3 Erklärung der Ausleuchtung

Abbildung 4-5 zeigt einen Ausschnitt des verwendeten Lupenbilds, auf dem alle – für die Durchführung der Studie notwendigen – Ausleuchtungen dargestellt sind. Die Abbildung zeigt exemplarisch den Originalbedienplatz der Stadtbahnsimulation. Die folgenden Erläuterungen gelten ebenfalls für den Experimentalbedienplatz.

In der Simulation gibt es zwei Arten von Signalen: Selbstblocksignale und Signale in direkter Verantwortung des Bedieners. Unterschieden werden diese Signale über die Signalbezeichnung. Eine gelbe Bezeichnung kennzeichnet ein Selbstblocksignal (S06) und eine rote bzw. grüne ein Signal in direkter Verantwortung (S09/S11). Grundsätzlich steht ein rotes Signal-symbol für ein haltzeigendes Signal (Haltbegriff) und ein grünes für ein fahrtzeigendes Signal (Fahrtbegriff). Zudem können über den Signalfuß viele weitere Informationen, wie z. B. die verzögerte D-Weg-Auflösung (blauer Fuß) oder die Fahrstraßenüberwachungsebene (gelber oder grüner Fuß) dargestellt werden.

Analog zu den Signalen besitzen auch die Gleisfreimeldeabschnitte verschiedene Ausleuchtungsvarianten. Eine gelbe Ausleuchtung des Abschnittes (Gleis oder Weiche) kennzeichnet die Grundstellung der Gleisfreimeldung. Grundstellung bedeutet in diesem Zusammenhang, dass der Gleisfreimeldeabschnitt frei und nicht belegt bzw. gestört ist. Eine grüne Ausleuchtung zeigt eine eingestellte Fahrstraße an und eine rote Ausleuchtung ist ein Zeichen für eine Belegung durch einen Triebzug oder eine Störung der Anlage. Über die Ausleuchtung des Gleisfreimeldeabschnitts der Weiche, wird zusätzlich die Weichenlage angezeigt. Die W04 in Abbildung 4-5 liegt in Linkslage. Der Fahrweg von S09 nach S20 ist eingestellt. Der nicht befahrbare Weichenschenkel ist weiß.

Eine Besonderheit der Weiche ist der Zuglenkauftragsmelder. Dabei handelt es sich um einen gelb blinkenden Richtungspfeil, der den Bediener über die für die nächste Zugfahrt notwendige Weichenlage informiert (W02 in Abbildung 4-5).

Die Zugnummer wird in einem weißen Rahmen neben dem dazugehörigen Gleisabschnitt angezeigt und besteht aus einer achtstelligen Kombination von Zahlen und Buchstaben. Über die Zugnummer werden dem Bediener Informationen über den Fahrweg und die Liniennummer des Triebzuges geben. In der Simulation gibt es eine Vielzahl weiterer Ausleuchtungen, auf deren Erklärung an dieser Stelle verzichtet wird, da sie für die Bedienung und das Verständnis der Bedienoberflächenstudie nicht relevant sind.

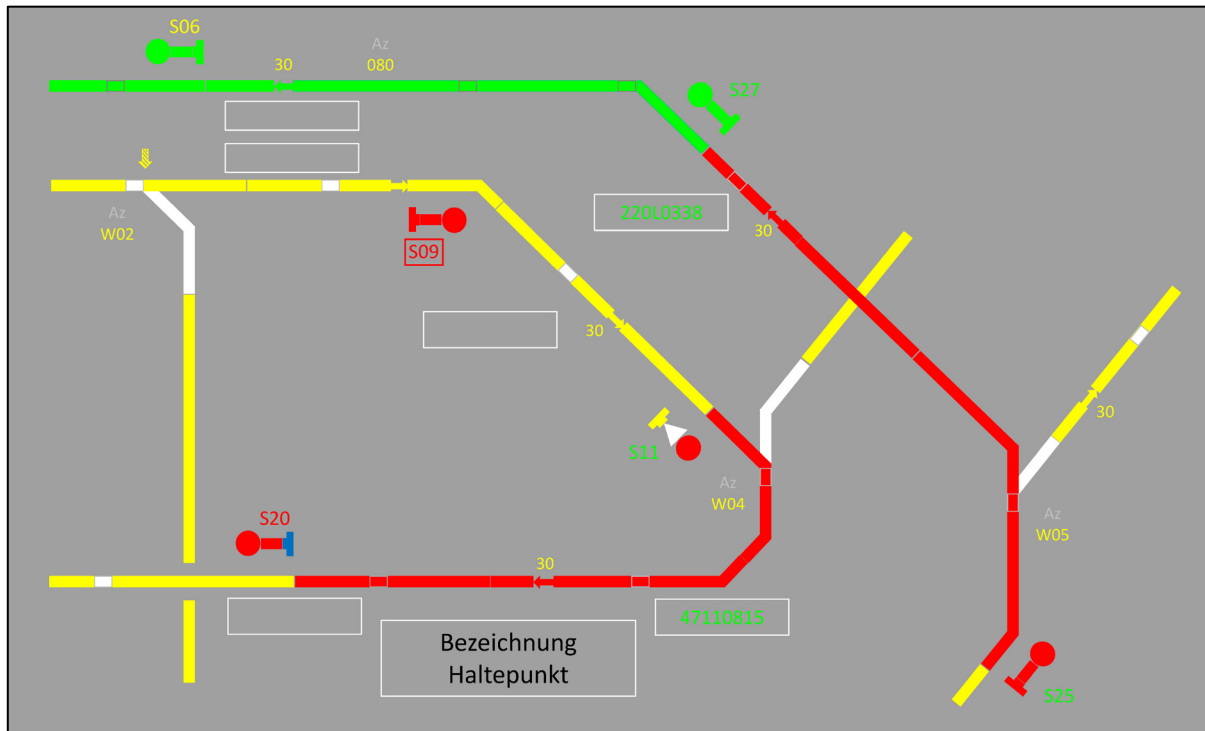


Abbildung 4-5: Ausschnitt des verwendeten Lupenbilds (Originalbedienplatz)

4.2.4 Polarität und Barrierefreiheit der Bedienplätze

In Kapitel 2.4 wurden die verschiedenen Darstellungsarten beschrieben. Im Zuge dessen wurde auf die Besonderheiten bei der Einteilung moderner Bedienoberflächen eingegangen und der Begriff *Polarität* eingeführt (Kapitel 2.4.2). In diesem Kapitel wird die Polarität der verschiedenen Bedienoberflächen bestimmt. Dafür wird vorab die relative Farbhelligkeit der verwendeten Farben benötigt. Hierfür muss der RGB-Wert der einzelnen Farben ermittelt werden. Für die RGB-Werte des Originalbedienplatzes konnte eine Systemdatei verwendet werden. Diese wurde dem Versuchsleiter⁴⁹ von der Firma Siemens zur Verfügung gestellt. Die RGB-Werte des Experimental- und des Designbedienplatzes wurden von dem Versuchsleiter vorgegeben und waren deshalb bekannt. Tabelle 4-1 zeigt eine Gegenüberstellung der Bedienplätze mit den verwendeten Farben und den dazugehörigen RGB-Werten. Die Übersicht beschränkt sich auf die wichtigsten und für die Rekonstruierbarkeit der Studie notwendigen Farben.

⁴⁹ Verfasser der vorliegenden Arbeit

Tabelle 4-1: RGB-Wert der einzelnen Bildelemente für Original-, Experimental- und Designbedienplatz

Bildschirmelement	Originalbedienplatz			Experimentalbedienplatz			Designbedienplatz		
	R	G	B	R	G	B	R	G	B
Bildschirm									
Hintergrund	160	160	160	0	0	0	180	180	180
Zugnummer	0	255	0	255	0	0	255	255	255
Signal									
rot	255	0	0	255	0	0	255	0	0
grün	0	255	0	34	139	34	34	139	34
gelb	255	255	0	255	255	0	255	255	0
blau	0	128	255	0	128	255	—	—	—
schwarz	—	—	—	—	—	—	0	0	0
Fahrweg									
rot	255	0	0	255	0	0	255	0	0
grün	0	255	0	34	139	34	34	139	34
gelb	255	255	0	255	255	0	255	255	0
weiß	255	255	255	255	255	255	255	255	255
dunkelgrau	0	128	255	—	—	—	90	90	90

4.2.4.1 Farbhelligkeit und Kontrast

Als Nächstes werden die relativen Farbhelligkeiten mithilfe der Formel 2-1 (Abbildung 2-12) aus Kapitel 2.3.2 berechnet. Tabelle 4-2 zeigt die Ergebnisse der Berechnung. Gerade bei dem Vergleich der verwendeten Farben fällt auf, dass die relativen Farbhelligkeiten des Originalbedienplatzes größer sind als die errechneten Farbhelligkeiten bei den anderen beiden Bedienplätzen. Dieser Umstand wird durch den Vergleich des Mittelwerts der relativen Farbhelligkeiten aller Bedienplätze nochmals verdeutlicht. Der Originalbedienplatz weist hier eine höhere Gesamthelligkeit auf.

Hinzu kommt, dass das mittlere Kontrastverhältnis des Originalbedienplatzes – im Vergleich zu den anderen beiden Bedienplätzen – den niedrigsten Wert hat. Der Grund hierfür ist die helle Hintergrundfarbe in Verbindung mit den grellen Vordergrundfarben. Die Gefahr dabei ist, dass Informationen durch das niedrige Kontrastverhältnis überstrahlt werden. Zudem wird das Erkennen von Unregelmäßigkeiten erschwert. Dieser Umstand wird im weiteren Verlauf der Studie noch näher betrachtet. Der Experimentalbedienplatz erreicht hier eindeutig bessere Werte, wodurch auch das Erkennen von Unregelmäßigkeiten erleichtert wird.⁵⁰

⁵⁰ Für den Nachweis wird die Fehlerentdeckungszeit ausgewertet und verglichen (Kapitel 3.1.5).

Tabelle 4-2: Übersicht über die relative Farbhelligkeit (FH) und das Kontrastverhältnis (KV)

Bildschirmelement	Originalbedienplatz		Experimentalbedienplatz		Designbedienplatz	
	FH	KV	FH	KV	FH	KV
Bildschirm						
Hintergrund	0.36	—	0.00	—	0.46	—
Zugnummer	0.72	1.87:1	0.21	5.25:1	1.00	2.04:1
Signal						
rot	0.21	1.56:1	0.21	5.25:1	0.21	1.96:1
grün	0.72	1.87:1	0.19	4.83:1	0.19	2.13:1
gelb	0.93	2.39:1	0.93	19.56:1	0.93	1.90:1
blau	0.23	1.46:1	0.23	5.58:1	—	—
schwarz	—	—	—	—	0.00	10.29:1
Fahrweg						
rot	0.21	1.56:1	0.21	5.25:1	0.21	1.96:1
grün	0.72	1.87:1	0.19	4.83:1	0.19	2.13:1
gelb	0.93	2.39:1	0.93	19.56:1	0.93	1.90:1
weiß	1.00	2.57:1	1.00	21.00:1	1.00	2.04:1
dunkelgrau	—	—	—	—	0.10	3.41:1
Mittelwert	0.60	1.95:1	0.41	10.12:1	0.48	2.98:1

4.2.4.2 Polarität der verwendeten Bedienoberflächen

Bei dem Experimentalbedienplatz sind die relativen Farbhelligkeiten aller Vordergrundfarben heller als die Hintergrundfarbe. Dies spricht für eine negative Polarität. Der Designbedienplatz verwendet, bis auf die sehr hellen Farben Weiß und Gelb, im Vergleich zur Hintergrundfarbe ausschließlich dunklere Vordergrundfarben. Er besitzt demnach eine positive Polarität. Der Originalbedienplatz verwendet hellere sowie dunklere Vordergrundfarben und wirkt insgesamt sehr inhomogen. Nichtsdestotrotz überwiegt der Anteil heller Vordergrundfarben, weshalb die Darstellungsart des Originalbedienplatzes eher der negativen Polarität entspricht.

4.2.4.3 Prüfung auf Barrierefreiheit

Aus [W3C09] kommen weitere Vorgaben, die sich mit dem Farbkonzept befassen. In Kapitel 2.3.2 wird gezeigt, dass im Allgemeinen das Farbkonzept des Bildschirms dem Anspruch der Barrierefreiheit genügen sollte. Diese Forderung gilt nur bedingt für den Fahrdienstleiterarbeitsplatz, da konzernrechtliche Bestimmungen u. a. die Farbenblindheit im Betriebsdienst nicht zulassen. Nichtsdestotrotz bieten die Vorgaben eine Möglichkeit, die verwendeten Bedienoberflächen auf ihre Barrierefreiheit und ihre Benutzerfreundlichkeit zu überprüfen. Tabelle 4-3 zeigt die Ergebnisse der Untersuchung. Neben der Helligkeitsdifferenz und dem Farbunterschied zwischen der Vorder- und der Hintergrundfarbe zeigt die dritte Spalte die

erreichte Konformitätsstufe nach den WCAG 2.0. Zur Erinnerung: Die Minstdifferenz der Farbhelligkeit (HD) sollte größer als 125 sein und der Farbunterschied (FU) zwischen beiden Farben sollte einen Mindestwert von 500 haben. Im Unterschied zur Textdarstellung werden die grundsätzlichen Informationen bei einer Bedienoberfläche über die farbliche Ausleuchtung der Fahrwegelemente dargestellt. Dies sind in der Regel größere Flächen, respektive Farbbalken. Dadurch ergibt sich eine deutlich höhere Erkennbarkeit, weshalb der Schriftgrad bei der Einteilung der Bedienoberfläche in die Konformitätsstufen (KS) eine untergeordnete Rolle spielt. Für die Einteilung in die KS wurde vorrangig das KV aus Tabelle 4-2 verwendet.

Die Prüfung auf Barrierefreiheit führt zu zwei Erkenntnissen. Zum einen zeigt die Untersuchung deutliche Unterschiede zwischen den Bedienoberflächen der verschiedenen Bedienplätze und zum anderen zeigt sich, dass bis auf wenige Ausnahmen die Vorgaben aus [W3C09] nur bedingt erfüllt werden. Gerade der Originalbedienplatz erfüllt in keinem Punkt die Vorgaben für ein barrierefreies interaktives System. Der Experimentalbedienplatz hingegen zeigt ein besseres Bild. Obwohl der Experimentalbedienplatz nur wenige Vorgaben in Bezug auf die HD und die FU erfüllt, erzielt er bei jeder Farbkonstellation eine sehr gute Konformitätsstufe. Der Designbedienplatz erfüllt ebenfalls nicht die Vorgaben aus [W3C09], erreicht im Vergleich zur hellen Bedienoberfläche jedoch eine bessere Bewertung.

Tabelle 4-3: Übersicht über die Helligkeitsdifferenz (HD) und den Farbunterschied (FU), die Werte in fett stehen für die Einhaltung der Vorgaben und die kursiven Werte für eine Zielverfehlung. Die dritte Spalte zeigt die erreichte Konformitätsstufe (KS) nach den WCAG 2.0.

Bildschirmelement	Originalbedienplatz			Experimentalbedienplatz			Designbedienplatz		
	HD	FU	KS	HD	FU	KS	HD	FU	KS
Bildschirm									
Hintergrund	—	—	—	—	—	—	—	—	—
Zugnummer	10	415	—	76	255	AAA	75	225	—
Signal			—						
rot	84	415	—	76	255	AAA	104	435	—
grün	10	415	—	96	207	AAA	84	333	—
gelb	66	350	—	226	510	AAA+	46	330	—
blau	56	287	—	104	383	AAA	—	—	—
schwarz	—	—	—	—	—	—	180	540	AAA+
Fahrweg									—
rot	84	415	—	76	255	AAA	104	435	—
grün	10	415	—	96	207	AAA	84	333	—
gelb	66	350	—	226	510	AAA+	46	330	—
weiß	95	285	—	255	765	AAA+	75	225	—
dunkelgrau	—	—	—	—	—	—	90	270	AA+
Mittelwert	54	372	—	137	372	—	89	346	—

Diese Ergebnisse zeigen, dass sich das Verfahren grundsätzlich eignet, um eine erste Aussage über die Benutzerfreundlichkeit der Bedienoberfläche zu treffen. Ob sich die gefundenen Unterschiede zwischen den Bedienoberflächen auf die Ergebnisse der Bedienoberflächenstudie auswirken, wird im weiteren Verlauf untersucht.

4.3 Betrieb in der Rückfallebene

In der Praxis gibt es eine Vielzahl von Unregelmäßigkeiten, die den Bahnbetrieb beeinflussen können. Eine Unregelmäßigkeit ist definiert als eine unzulässige Abweichung der Signalanlage vom Sollzustand [DBAG13b]. Sie wird in zwei Arten unterschieden: Fehler und Störung. Ein Fehler ist eine Unregelmäßigkeit, die den Betrieb nicht beeinflusst. Beispielsweise wird der Ausfall des Hauptfadens einer Zweifadensignallampe dem Bediener in Form eines „Blinkers“ angezeigt. Beim ESTW blinkt der Signalschirm auf dem Lupenbild. Der Betrieb wird dadurch jedoch nicht eingeschränkt. Fällt im vorangegangenen Beispiel der Nebenfaden zusätzlich aus, kommt es zu einer Störung. In diesem Fall kann das Signal erloschen sein. Demnach ist eine Störung eine Unregelmäßigkeit, die den Betrieb beeinflusst.⁵¹

Der Einfachheit halber werden die Versuchsteilnehmer nur mit drei verschiedenen Unregelmäßigkeiten konfrontiert. Diese setzen sich aus einem Fehler und zwei Störungen zusammen. Der Fehler in der Simulation ist schnell behoben, bei den Störungen muss der Bediener eine Reihe von registrierungspflichtigen Hilfshandlungen durchführen.

4.3.1 Hilfshandlungen und Kommandofreigabe

Durch registrierungspflichtige Hilfshandlungen kann die betriebliche Verfügbarkeit aufrechterhalten werden. In der Stadtbahnsimulation werden Hilfshandlungen mittels Kommandofreigabeverfahren durchgeführt. Das Kommandofreigabeverfahren dient der Sicherung von sicherheitsrelevanten Bedieneingaben [FoWe94] und zählt zu der technischen Verfahrenssicherung. Mit dem Verfahren sollen u. a. zwei Eingabefehlerarten verhindert werden. Zum einen die Eingabefehler, die auf Denkfehlern beruhen und zum anderen die, die durch ein falsches Meldebild entstehen [Wie07].

Die Hilfshandlung wird durch das Anwählen des zu bedienenden Elements und die Auswahl der jeweiligen Hilfshandlung eingeleitet. Dadurch wird der eingegebene Befehl an die Sicherungsebene übertragen und die Anzeigensicherung initiiert. Der Bediener wird durch ein Dialogfenster aufgefordert den ausgegebenen Befehl mit dem eingegebenen zu vergleichen. Hierfür wird ein eigenes Dialogfenster am linken unteren Bildschirmrand aufgeschaltet (Abbildung 4-6).

⁵¹ In der vorliegenden Arbeit werden die Begriffe *Unregelmäßigkeit*, *Störung* und *Fehler* synonym verwendet. Jeder dieser Begriffe steht für einen ungewollten Zustand.



Abbildung 4-6: Dialogfenster für die Kommandofreigabe

Der Bediener muss prüfen, ob laut Meldebild die Freigabe betrieblich zulässig ist und ob Farbbalken und Aktualisierungsmelder korrekt dargestellt werden. Dabei muss sich der Aktualisierungsmelder (Kreis mit rotem Balken) drehen. Durch die Drehbewegung erzwingt der Aktualisierungsmelder einen konstanten Prozessdatenfluss und versorgt die Anzeige dadurch mit aktuellen Daten. Die zweizeiligen Farbbalken sichern die korrekte Farbdarstellung. Die obere Reihe ist statisch und die untere blinkt. Ist dies der Fall, betätigt der Bediener die erste Kommandofreigabe (KF1) und gibt die Hilfshandlung frei. Jetzt bleiben ihm 15 s, um eine erneute Bildschirmprüfung durchzuführen und die Eingabe der Hilfshandlung zu überprüfen. Hat er sich nochmals vergewissert, bestätigt er im Anschluss die zweite Kommandofreigabe (KF2). Die Hilfshandlung wird ausgeführt und automatisch protokolliert. Dieser Ablauf ist bei jedem registrierungspflichtigen Kommando notwendig.

4.3.2 Unregelmäßigkeiten während des Versuchs

Die vorgesehenen Unregelmäßigkeiten unterscheiden sich sehr stark in der betrieblichen Einschränkung und im Umfang der notwendigen Hilfshandlungen. Die Schwierigkeit und die Abfolge der Unregelmäßigkeiten werden so gewählt, dass die Versuchsteilnehmer gefordert, aber nicht überfordert werden. Zudem dient die unterschiedliche Schwierigkeit dazu, den Wissensstand der Versuchsteilnehmer abzufragen, die Trainingsdefizite aufzuzeigen sowie unnötige Bedienhandlungen schneller zu erkennen. Des Weiteren ist für den Umgang mit den Unregelmäßigkeiten nur ein begrenztes betriebliches Wissen notwendig und sie können gut nach einem Schema erlernt werden. Es müssen keine betrieblichen Maßnahmen eingeleitet werden und es sind keine technischen Vorbedingungen für den Eingabebefehl zu erfüllen.

Nachfolgend wird der Umgang mit den Störungen erläutert und die Anzahl der registrierungspflichtigen Hilfshandlungen aufgezeigt.

4.3.2.1 Ausfall Zuglenkbetrieb

Der Ausfall des Zuglenkbetriebs ist ein Fehler. Die Unregelmäßigkeit hat theoretisch keine Auswirkung auf den Betrieb. Der ZLB, respektive die Zuglenkung, ist eine Automatisierungsanlage zur selbsttätigen Einstellung von Zugfahrstraßen im Stellwerk. Das Stellwerk identifiziert die Züge anhand ihrer Zugnummer und stellt die jeweilige Fahrstraße ein. Die Funktion

des ZLB wird über die Signalbezeichnung abgebildet. Bei eingeschaltetem ZLB wird die Signalbezeichnung in grün dargestellt. Ist die Signalbezeichnung hingegen rot, bedeutet das, dass der ZLB ausgeschaltet ist (Abbildung 4-7).

Fällt der ZLB während der Simulation aus, funktioniert die zugbewirkte Fahrstraßenwahl nicht. Im Idealfall erkennt der Bediener den Ausfall direkt, schaltet den ZLB wieder ein und es kommt zu keinen Einschränkungen bzw. zu keinen Verspätungen. Erkennt er den Ausfall nicht sofort oder kann er die Funktion aufgrund eines Anlagendefekts nicht wieder einschalten, muss er die Fahrstraßen manuell einstellen. Durch die hohe Zugfolge im Stadtbahnnetz kann es hierdurch zu Verspätungen kommen. Da es sich bei dem Ausfall des Zuglenkbetriebs um einen Fehler handelt, sind keine Hilfshandlungen notwendig.

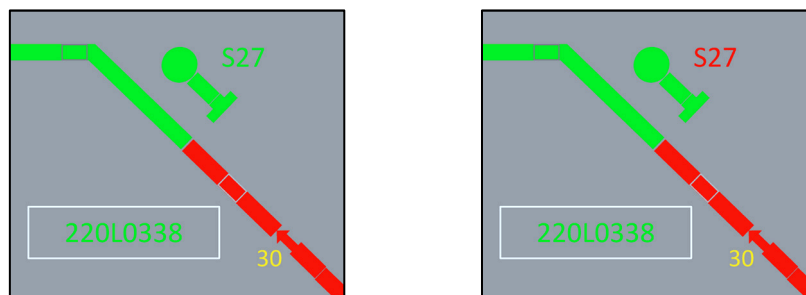


Abbildung 4-7: Darstellung Ausfall Zuglenkbetrieb S27 (links: ZLB eingeschaltet/rechts: ZLB ausgeschaltet)

4.3.2.2 Signalstörung (Grünlampendefekt)

Die zweite Unregelmäßigkeit ist die Signalstörung. Hierbei handelt es sich – wie der Name bereits erkennen lässt – um eine Störung. In der Bedienoberflächenstudie fällt die Grünlampe des Signals komplett aus. Auf dem Lupenbild blinkt der Signalkopf und eine Fahrtstellung des betroffenen Signals ist nicht mehr möglich (Abbildung 4-8).

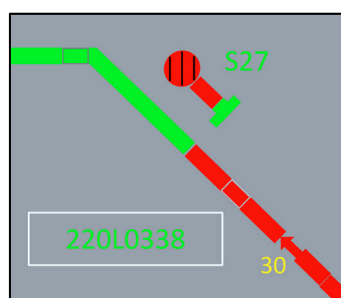


Abbildung 4-8: Darstellung Signalstörung S27 (Signalkopf blinkt)

Bevor der Bediener eine Hilfshandlung an dem gestörten Signal durchführen kann, muss er das betroffene Signal sperren. Die Signalsperrung ist eine notwendige Bedingung und wird von der Anlage überprüft. Für die Sperrung des Signals ist keine Hilfshandlung notwendig, da dieses Bedienkommando zur sicheren Seite geht. Ist das Signal nicht gestört, führt die Sperrung nur zu einer Einschränkung der Leistungsfähigkeit, da ein gesperrtes Signal nicht auto-

matisch auf Fahrt geht, auch wenn der ZLB eingeschaltet ist. Ist die Sperrung erfolgreich, wird ein roter Kasten um die Signalbezeichnung angezeigt (S09 in Abbildung 4-5). Der nächste Schritt ist das Stellen des Vorsichtssignals. Hierdurch erlaubt der Bediener dem Zug die Vorbeifahrt am gestörten Signal mit verringerter Geschwindigkeit. Das Vorsichtssignal wird in Form eines weißen Dreiecks unter dem Signalkopf dargestellt (S11 in Abbildung 4-5).

In der Regel treten der Ausfall des ZLB und die Signalstörung getrennt voneinander auf. Das heißt, der Bediener muss die Fahrstraße nicht händisch einstellen, da dies durch den ZLB erfolgt. Kommt es unerwartet trotzdem vor, muss der Bediener, bevor er das Vorsichtssignal stellen kann, die Fahrstraße händisch einstellen.

Die Signalsperrung bleibt bis zur Behebung der Störung erhalten. Nach Entstörung der Anlage – durch den Versuchsleiter – hört der Signalkopf wieder auf zu blinken. Der Bediener muss das erkennen und das Signal entsperren. Das Entsperren ist sicherheitsrelevant und wird deshalb als registrierungspflichtige Hilfshandlung gezählt. Die Art und Anzahl der notwendigen Hilfshandlungen wird in Tabelle 4-4 gezeigt.

Tabelle 4-4: Übersicht der Hilfshandlungen bei einer Signalstörung

Hilfshandlung	Bezeichnung	Anzahl
VS	Vorsichtssignal Stellen	1
SES	Signal EntSperren	1
Anzahl registrierungspflichtiger Hilfshandlungen: 2		

4.3.2.3 Gleisfreimeldestörung einer Weiche

Bei der dritten Unregelmäßigkeit zeigt der Gleisfreimeldeabschnitt eine Belegung an, obwohl der Triebzug den Abschnitt bereits geräumt hat. In der Abbildung 4-9 zeigt der Gleisfreimeldeabschnitt der Weiche 4 fälschlicherweise eine Belegung an.

Die Gleisfreimeldestörung tritt in der Studie immer nach einer Zugfahrt auf und ist am schwersten zu erkennen. Sie kann nur erkannt werden, wenn der Versuchsteilnehmer alle Zugfahrten beim Durchqueren des Bedienbereichs beobachtet oder die Betriebssituation richtig einschätzt und „Staus“ erkennt. Hinzu kommt, dass der Umfang der notwendigen Hilfshandlungen bei der Gleisfreimeldestörung am größten ist. Aus diesem Grund beansprucht diese Art von Störung mehr geistige Ressourcen. Es wird vermutet, dass ein ungeübter Teilnehmer durch diese Art von Störung eher aus der Ruhe gebracht wird, als durch die anderen beiden Unregelmäßigkeiten.

Um den Betrieb aufrechtzuerhalten zu können, muss der Bediener eine Reihe von Bedienhandlungen durchführen. Im Regelbetrieb erfolgen das Markieren und das Umstellen der Weichen automatisch, d. h. zugbewirkt. Im Falle einer Gleisfreimeldestörung muss der Bediener die betroffenen Gleisfreimeldeabschnitte (Weiche und Gleisabschnitt) mittels Hilfsauflösung

selbstständig in Grundstellung bringen. Der Grund hierfür ist, dass die Anlage alle weiteren Fahrwegelemente in der Fahrstraße nicht mehr automatisch auflöst. Das bedeutet, dass ab dem gestörten Element die darauffolgenden Fahrwegelemente mit einer registrierungspflichtigen Hilfshandlung aufgelöst werden müssen. Ob ein Fahrwegelement noch verschlossen ist, wird dem Bediener über den Verschlussmelder angezeigt. Dieser Melder wird als kleiner Block im Fahrstraßenband dargestellt. Im Beispiel wird der Verschlussmelder des gestörten Gleisfreimeldeabschnitts der Weiche rot ausgeleuchtet (Abbildung 4-9).



Abbildung 4-9: Darstellung Gleisfreimeldestörung W04 (mit hervorgehobenem Verschlussmelder an der Weiche und am darauffolgenden Gleisabschnitt)

Sind die Verschlussmelder aller Elemente in Grundstellung, muss der Bediener als nächstes die Fahrstraße selbstständig einstellen. Liegt eine gestörte Weiche im Fahrweg, muss er diese ggf. in die richtige Lage bringen. Das macht der Bediener mittels Weichenumstellung bzw. bei gestörtem Gleisfreimeldeabschnitt der Weiche mittels Weichenhilfsumstellung. Letzteres ist eine registrierungspflichtige Handlung. Wenn alle Elemente sich in der für die Fahrstraße richtigen Lage befinden, kann die Fahrstraße eingestellt werden. Anschließend erlaubt der Bediener die Vorbeifahrt am Signal mit dem Vorsichtssignal. Die Art und Anzahl der notwendigen Hilfshandlungen ist abhängig von der Lage der Störung und der darauffolgenden Zugfahrt (Tabelle 4-5).

Tabelle 4-5: Übersicht der Hilfshandlungen bei einer Gleisfreimeldestörung an einer Weiche

Hilfshandlung	Bezeichnung	Anzahl
WHA	Weiche HilfsAuflösen	1
GAHA	GleisAbschnitt HilfsAuflösen	min. 1
WHU	Weiche HilfsUmstellen	ggf.
VS	Vorsichtssignal Stellen	1
Anzahl registrierungspflichtiger Hilfshandlungen: min. 3		

4.3.3 Fehlerstatusanzeige

An dieser Stelle ist noch wichtig zu erwähnen, dass Störungen der Anlage dem Bediener über eine Statusanzeige im oberen Bereich des Bildschirms angezeigt werden (Abbildung 4-10). Dabei ist die Darstellung der Fehlerstatusanzeige unabhängig vom verwendeten Bedienplatz. Die Buttons *STW-Auswahl* und *ALARM* blinken bei einer Anlagenstörung rot. Durch Anklicken des Buttons *STW-Auswahl* wird dem Bediener ein Auswahlfenster gezeigt, auf dem alle Stellwerke inkl. aller einstellbaren Lupenbilder aufgelistet sind. Die Bezeichnung der Lupenbilder, auf denen eine Störung ansteht, blinkt rot. Durch Anklicken des Buttons *ALARM* verschwindet das Lupenbild und der Bediener erhält eine Übersicht aller anstehenden Störungen. Diese muss er zur Kenntnis nehmen und quittieren.

Ein Vorteil der Fehlerstatusanzeige ist, dass der Bediener direkt erkennt, dass in seinem Zuständigkeitsbereich etwas nicht in Ordnung ist. Demgegenüber steht ein für die Studie nicht unerheblicher Nachteil. Die Buttons der Fehlerstatusanzeige hören erst auf zu blinken, wenn alle Störungen auf allen Lupenbildern, quittiert sind. In der Simulation kann es vorkommen, dass auf einem anderen Lupenbild Triebzüge ungewollt fehlgeleitet werden. Dadurch bekommen sie eine Fehlernummer als Zugnummer, was der Anlage gemeldet wird und als Störung angezeigt wird. Als mögliche Folge lösen Fahrstraßen nicht ordnungsgemäß auf, wodurch Hilfshandlungen auf anderen Lupenbildern notwendig werden. Hierfür muss der Bediener das Lupenbild wechseln. Um die Fehlerentdeckungszeit der Versuchsteilnehmer nicht zu verfälschen, wird der Lupenbildwechsel in der Studie nicht geschult.

Um die Versuchsteilnehmer durch die rot blinkenden Alarme nicht zu irritieren, wird die Fehlerstatusleiste im Experiment abgedeckt.⁵² Hierdurch entstehen keine Nachteile für die Versuchsteilnehmer. Sie werden eher entlastet und es werden gleiche Voraussetzungen für alle Teilnehmer geschaffen.

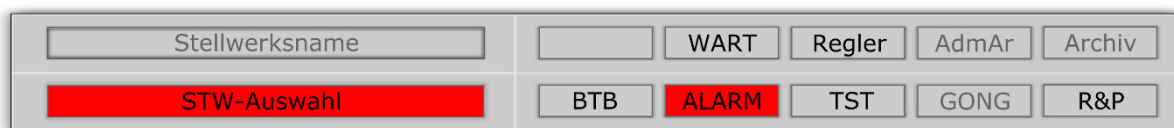


Abbildung 4-10: Darstellung Fehlerstatusanzeige

Auf die Erklärung der anderen Buttons wird an dieser Stelle verzichtet. Sie sind für die Bedienung und das Verständnis der Bedienoberflächenstudie nicht relevant.

⁵² siehe hierzu Abbildung 5-7

5 BEDIENOBERFLÄCHENSTUDIE

Um den Einfluss der farblichen Gestaltung der Bedienoberfläche auf den Bediener zu untersuchen, wurden im Rahmen einer Studie zwei Bedienoberflächen mit negativer Polarität verglichen. Die untersuchten Bedienoberflächen unterschieden sich nur hinsichtlich der farblichen Gestaltung, nicht aber in Bezug auf Benutzbarkeit und Funktionalität.

Ziel dieses Kapitels ist die Vorstellung der Bedienoberflächenstudie. Das Kapitel gliedert sich in drei Teile. Im ersten Teil wird die Methode erläutert (Kapitel 5.1). Dafür wird zunächst die Stichprobe beschrieben, im Anschluss daran die verwendeten Materialien und Geräte vorgestellt, die abhängigen Variablen operationalisiert und zum Schluss die Versuchsdurchführung skizziert. Im zweiten Teil dieses Kapitels werden die Ergebnisse vorgestellt und statistisch bewertet (Kapitel 5.2). Abschließend werden die Ergebnisse diskutiert (Kapitel 5.3).

5.1 Methode

Der Studie lag ein unverbundenes Zweistichprobenproblem zugrunde. Dabei wurden die Faktoren *farbliche Gestaltung* und *Geschlecht* untersucht. Beide Faktoren lagen in zwei Stufen vor (Kapitel 5.1.2.2).

Zur Untersuchung wurden zwei unabhängige Versuchsgruppen gebildet. Die Gruppeneinteilung erfolgte nach dem *FIFO-Prinzip* (First In–First Out). Dies ermöglichte eine zufällige Verteilung der Versuchsteilnehmer. Der Vorteil dieses Vorgehens ist, dass mögliche Störvariablen weder bekannt noch operationalisiert oder gemessen werden müssen [Hub10], wodurch die Versuchsdurchführung vereinfacht wird. Da beide Bedienplätze eine Bedienoberfläche mit negativer Polarität besitzen, werden in diesem Kapitel neue Begrifflichkeiten eingeführt, um die Unterscheidung zu erleichtern. Für die Bedienoberfläche des Originalbedienplatzes wird im Fließtext der Begriff *helle Bedienoberfläche* und in Diagrammen die Bezeichnung *Hell* verwendet. Bei dem Experimentalbedienplatz wird der Begriff *dunkle Bedienoberfläche* im Fließtext und die Bezeichnung *Dunkel* in Diagrammen eingeführt.

Bevor die Bedienoberflächenstudie im V-EBL durchgeführt wurde, fand im Rahmen der Lehrveranstaltung *Projekte des Bauingenieurwesens* im Sommersemester 2015 eine Voruntersuchung statt. Dabei wurden der Original-, der Experimental- und der Designbedienplatz an $N = 10$ Versuchsteilnehmern getestet, um erste Erfahrungen zu sammeln und verschiedene

Fragebögen auf ihre Anwendbarkeit zu überprüfen. Die Ergebnisse dieses Tests flossen in die folgenden Beschreibungen mit ein.

5.1.1 Stichprobe

Insgesamt nahmen $N = 52$ Versuchsteilnehmer an der Studie teil. Die Gruppe der Versuchsteilnehmer setzte sich aus 45 Studierenden verschiedener Fachrichtungen, vier Angestellten der Universität und drei Privatpersonen zusammen. Die Teilnehmer wurden über Bekanntgaben in Vorlesungen, auf Facebook und über Stud.IP⁵³ angeworben. Die 25 weiblichen und 27 männlichen Versuchsteilnehmer waren zwischen 15 und 54 Jahre alt. Das Durchschnittsalter betrug 25.23 Jahre ($SD^{54} = 6.78$).

Abbildung 5-1 zeigt die Zusammenstellung der beiden Versuchsgruppen. In der Gruppe mit der hellen Bedienoberfläche befanden sich 13 weibliche und 16 männliche Versuchsteilnehmer im Alter von 19 bis 54 Jahren ($M^{55} = 26.21$, $SD = 8.27$). Die Versuchsgruppe mit der dunklen Bedienoberfläche setzte sich aus 12 weiblichen und 11 männlichen Versuchsteilnehmern zusammen. Diese waren zwischen 15 und 32 Jahre alt und das Durchschnittsalter betrug 24.00 Jahre ($SD = 4.09$).

Keiner der Teilnehmer litt an einer Farbsehschwäche⁵⁶ und verfügte über Vorerfahrungen mit der Versuchsumgebung bzw. mit der Simulation. Für ihre Teilnahme erhielten die Versuchsteilnehmer eine Aufwandsentschädigung von 30 Euro.

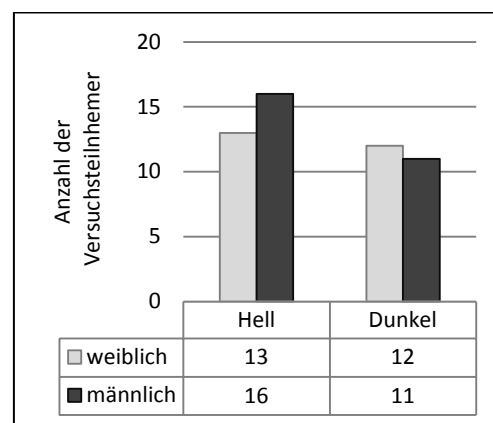


Abbildung 5-1: Aufteilung der Teilnehmer auf die Gruppen und nach Geschlecht

⁵³ Stud.IP (**Studien**begleitender Internetsupport von **Präsenz**lehre) ist eine internetbasierte Arbeitsumgebung zur Unterstützung der Lehre.

⁵⁴ SD = Standardabweichung (*standard deviation*)

⁵⁵ M = Mittelwert (*mean*)

⁵⁶ Hierfür wurden die *Tafeln zur Prüfung des Farbsinnes* verwendet [BKS11].

5.1.2 Versuchsplan

Wie in Kapitel 2 gezeigt wurde, hat die farbliche Gestaltung der Bedienoberfläche einen nicht unerheblichen Einfluss auf verschiedene Faktoren, wie zum Beispiel die subjektive Wahrnehmung der Usability. Im Eisenbahnwesen gibt es bisher keine Untersuchung, die den Einfluss der farblichen Gestaltung der Bedienoberfläche auf den Stellwerksbediener betrachtet. Aus diesem Grund wurde eine explorative Studie durchgeführt, in der verschiedene Ansätze gewählt wurden, um den Einfluss auf die Versuchsteilnehmer zu untersuchen.

5.1.2.1 Hypothesen

Um die Orientierung in der Bedienoberflächenstudie zu erleichtern, wurden sechs Hypothesen formuliert. Diese stützten sich auf Erkenntnisse aus anderen Industriezweigen und auf die qualitativen Ergebnisse der Voruntersuchung. Durch die Hypothesen werden die Erkenntnisse auf das Eisenbahnwesen übertragen und mit der Zielsetzung aus Kapitel 3 verknüpft. Zudem werden die Tendenzen aus der Voruntersuchung mithilfe einer größeren Stichprobe auf ihre Gültigkeit überprüft.

Die erste Hypothese befasst sich mit der Untersuchung des Einflusses der farblichen Gestaltung. Da die Richtung des Einflusses nicht bekannt war, wurde sie ungerichtet formuliert:

Hypothese 1 (ungerichtet): Die farbliche Gestaltung der Bedienoberfläche beeinflusst die wahrgenommene User Experience, die empfundene psychische Beanspruchung, die Befindlichkeit, die subjektive Wahrnehmung der Bedienungssicherheit und die erfassten Leistungskennzahlen.

Um die Richtung eines möglichen Einflusses untersuchen zu können, wurden zwei weitere gerichtete Hypothesen formuliert:

Hypothese 2 (gerichtet): Die User Experience, die psychische Beanspruchung und die Befindlichkeit der Versuchsteilnehmer haben einen Einfluss auf die Beurteilung der Bedienungssicherheit. Im Einzelnen wird davon ausgegangen, dass sich hohe Werte bei der wahrgenommenen User Experience, eine geringe Beanspruchung während des Erlernens und eine gute Befindlichkeit nach der Eingewöhnungsphase positiv auf die Bedienungssicherheit auswirken.

Hypothese 3 (gerichtet): Eine hohe Bewertung der User Experience und der Bedienungssicherheit sowie eine gute Befindlichkeit, schlagen sich in einer besseren Bedienereffizienz nieder.

Dabei hat die Bedienungssicherheit einen wesentlichen Einfluss auf das Überwachungsverhalten der Versuchsteilnehmer. Die Annahme ist, dass Versuchsteilnehmer, die sich bei der Bedienung sicher fühlen, zu jeder Zeit ein besseres Situationsbewusstsein besitzen als unsichere Teilnehmer. Daher fällt es ihnen leichter die Betriebssituation trotz Ablenkung schnell

und richtig einzuschätzen. Sie haben mehr Reservekapazität (Kapitel 5.1.3.10) zur Verfügung und können deshalb der Nebenaufgabe eine höhere Aufmerksamkeit widmen, ohne dabei einen Leistungseinbruch bei der Überwachung des Bedienbereichs zu riskieren. Der letztgenannte Aspekt ist deshalb von Interesse, da der Versuchsleiter die Aufmerksamkeit, die der Nebenaufgabe gewidmet wird, nicht steuern kann. Möglicherweise dominiert die Nebenaufgabe die Hauptaufgabe. Wäre dies der Fall, würde die Leistung in der Hauptaufgabe unbewusst negativ beeinflusst werden [WHB+13]. Diesen Umständen wurde mit folgender Hypothese Rechnung getragen:

Hypothese 4 (gerichtet): Versuchsteilnehmer, die sich sicher fühlen, d. h. ihre Bedienungssicherheit selbst hoch einschätzen, verwenden weniger Zeit mit der Überwachung des Bedienbereichs und erzielen ein besseres Ergebnis in der Nebenaufgabe, wobei sich die einzelnen Fehlerentdeckungszeiten durch die geringere Überwachung nicht signifikant erhöhen.

Im nächsten Schritt sollen die qualitativen Ergebnisse der Voruntersuchung mithilfe der Bedienoberflächenstudie überprüft und verifiziert werden. Eine Erkenntnis der Voruntersuchung war, dass die Versuchsteilnehmer die dunkle Bedienoberfläche favorisierten. Dabei wurden Eigenschaften wie Übersichtlichkeit, Kontrastwirkung und Wahrscheinlichkeit der Fehlererkennung besser bewertet als bei der hellen Bedienoberfläche. Hieraus ergibt sich folgende gerichtete Hypothese:

Hypothese 5 (gerichtet): Die dunkle Bedienoberfläche erhält, im Vergleich zur hellen Bedienoberfläche, eine bessere Bewertung bei der wahrgenommenen User Experience. Sie wird als attraktiver, gebrauchstauglicher und stimulierender empfunden.

Hierbei ist vor allem die stimulierende Wirkung der Bedienoberfläche von Interesse. Es wird vermutet, dass durch die aktivierende Wirkung der farblichen Gestaltung die Versuchsteilnehmer weniger Zeit benötigen, um eine Unregelmäßigkeit zu erkennen. Diese Annahme mündet in der folgenden gerichteten Hypothese:

Hypothese 6 (gerichtet): Die Versuchsteilnehmer in der Gruppe mit der dunklen Bedienoberfläche sind aktiver und benötigen dadurch weniger Zeit, um eine Unregelmäßigkeit zu erkennen.

5.1.2.2 Unabhängige Variablen

Wie bereits beschrieben, wurden die Einflussfaktoren *farbliche Gestaltung* und *Geschlecht* untersucht. Dafür wurden zwei unterschiedliche Bedienoberflächen miteinander verglichen. Da sich diese ausschließlich in der farblichen Gestaltung voneinander unterschieden, lag auch das Hauptaugenmerk der Studie auf der Untersuchung des Einflusses der farblichen Gestaltung. Das Geschlecht der Versuchsteilnehmer wurde zusätzlich als Faktor untersucht, um mögliche Einflüsse des Geschlechts auf die Ergebnisse der Bedienoberflächenstudie zu untersuchen. Durch dieses Vorgehen wurde das Geschlecht als mögliche Störvariable kontrolliert. Weitere Faktoren wie das Studienfach, persönliche Interessen und Fähigkeiten oder

das Motiv für die Teilnahme waren denkbar. Diese wurden aus versuchsökonomischen Gründen nicht betrachtet. Es handelte sich rein um geschlechterspezifische Unterschiede bei der Bewertung und Benutzung der Bedienoberfläche. Zudem war die Untersuchung des Geschlechts naheliegend, da die Daten ohne zusätzlichen Aufwand bereits zur Verfügung standen. Ebenso hätte das Alter der Versuchsteilnehmer als weitere Einflussvariable verwendet werden können. Hierauf wurde jedoch bewusst verzichtet, da der Großteil der Teilnehmer zwischen 20 und 29 Jahre alt war. Das entspricht einer Altersgruppe, weshalb eine weitere Clusterung – aus Sicht des Verfassers – nicht sinnvoll erschien. Demnach ergaben sich zwei unabhängige Variablen, welche in der Studie betrachtet wurden:

- die farbliche Gestaltung der Bedienoberfläche (zwei Stufen) und
- das Geschlecht der Versuchsteilnehmer (zwei Stufen).

5.1.2.3 Abhängige Variablen

Um die Hypothesen aus Kapitel 5.1.2.1 zu untersuchen und den Einfluss der farblichen Gestaltung der Bedienoberfläche auf den Bediener zu ermitteln, wurden folgende abhängige Variablen definiert:

- User Experience,
- psychische Beanspruchung,
- Befindlichkeit,
- Bedienungssicherheit und
- vordefinierte Leistungskennzahlen.

Die abhängigen Variablen wurden mit verschiedenen Messmitteln erfasst. Hierbei kamen hauptsächlich Fragebögen zum Einsatz – wie das nachfolgende Kapitel 5.1.3 zeigt. Darüber hinaus wurden weitere Gerätschaften für die Erfassung der Leistungskennzahlen verwendet. Dies war notwendig, um eine Aussage über die Bedienereffizienz treffen zu können. In Kapitel 5.1.4 werden das zugehörige Modell und die Leistungskennzahlen im Detail beschrieben.

5.1.3 Material und Geräte

Zur Untersuchung der in Kapitel 3 aufgestellten Zielstellung, wurde eine Reihe von Hilfsmitteln verwendet. Diese werden im nachfolgenden Kapitel vorgestellt und im Detail erklärt.

5.1.3.1 Hardware des Bedienplatzes

Bei der Hardware des Bedienplatzes handelte es sich um einen Standard-Windows-Heimcomputer (Intel Core i3, Windows 7 in der 64 Bit-Version) mit einer ATI Radeon HD 5450 Grafikkarte. Als Ausgabegerät wurde ein flimmerfreier BenQ GL2450 LED-Monitor mit einem dynamischen Kontrastverhältnis von 12.000.000:1 verwendet. Der Monitor wurde

über HDMI mit dem Rechner verbunden und befand sich in der Werkseinstellung. Für die Nebenaufgabe wurde ein Lenovo ThinkPad T400 (Intel Core 2 Duo, Windows 7 in der 64 Bit-Version) verwendet.

5.1.3.2 Fragebogen zur Beurteilung der User Experience

Wie bereits in Kapitel 2.2.2.2 beschrieben, ist die UX eine Erweiterung der klassischen Usability. Dadurch sind die UX-Fragebögen so konstruiert, dass sie auch Teile der Usability erfassen. Aus diesem Grund wurde in der Bedienoberflächenstudie ein UX-Fragebogen verwendet, um gleichzeitig eine Aussage über die UX und die Usability treffen zu können. Das straffte den Versuchsablauf und die Versuchsteilnehmer wurden nicht unnötig durch weitere Fragebögen belastet. Die gängigsten Fragebögen für die Bewertung der UX sind der AttrakDiff2 und der UEQ (Kapitel 2.2.2.2.3). Wie alle übrigen verwendeten Fragebögen wurden der AttrakDiff2 und der UEQ in der Voruntersuchung auf ihre Anwendbarkeit überprüft. In der Tendenz kamen beide Fragebögen zu ähnlichen Ergebnissen – siehe exemplarisch [LSH06].

Der entscheidende Nachteil des AttrakDiff2 war die webbasierte Befragung und Auswertung. Die Laborrechner besaßen keinen Internetzugang, wodurch ein neuer, papiergebundener Fragebogen notwendig gewesen wäre, um die Daten zu erfassen und später webbasiert auswerten zu können. Aus diesem Grund wurde der UEQ für die Beurteilung der UX verwendet (Anhang C1).

Der UEQ besteht aus 26 bipolaren Items. Diese Items sind auf die sechs Dimensionen *Attraktivität*, *Durchschaubarkeit*, *Effizienz*, *Steuerbarkeit*, *Stimulation* und *Originalität* verteilt. Als Antwortformat wird eine Rating-Skala mit 7 Stufen verwendet (Abbildung 5-2).⁵⁷ Ein Vorteil des UEQ ist, dass die sechs Dimensionen auf das theoretische Rahmenkonzept von *HASSENZAHN* (Kapitel 2.2.2.2) übertragen werden können. Hierfür werden die Einzelmittelwerte der Dimensionen logisch zusammengefasst, um die pragmatische Qualität (PQ), die hedonische Qualität (HQ) und die Attraktivität der Bedienoberflächen zu erhalten. Die pragmatische Qualität ergibt sich aus den Mittelwerten für Durchschaubarkeit, Effizienz und Steuerbarkeit und die hedonische Qualität setzt sich aus den Mittelwerten für Stimulation und Originalität zusammen. Die Attraktivität gilt als globale Dimension und kann direkt ausgewertet werden.

	1	2	3	4	5	6	7	
unerfreulich	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	erfreulich
unverständlich	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	verständlich
kreativ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	phantasielos
leicht zu lernen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	schwer zu lernen
wertvoll	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	minderwertig
langweilig	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	interessant

Abbildung 5-2: Auszug aus dem UEQ

⁵⁷ Für eine umfassende Beschreibung des Fragebogens wird auf [RTS13] verwiesen.

5.1.3.3 Fragebogen zur Beurteilung der psychischen Beanspruchung

Sind die Anforderungen, welche an eine Person gestellt werden, größer als die Leistungsvoraussetzung der Person, führt dies zu einer Überforderung [RiHa98]. Mit anderen Worten, auf eine zu hohe Belastung folgt eine psychische Beanspruchung. Dabei beschreibt die Belastung die Summe der objektiv messbaren Größen, die auf den Bediener einwirken und die Beanspruchung spiegelt die subjektive Wahrnehmung dieser wider [Ric00]. Das Problem daran ist, dass die psychische Beanspruchung einen Einfluss auf die Zuverlässigkeit und die Arbeitsweise des Bedieners hat, was wiederum die Sicherheit und die Wirtschaftlichkeit des Bahnbetriebs beeinflussen kann. Aus diesem Grund wurde die Beanspruchung der Versuchsteilnehmer während des Erlernens der Simulation erfasst, um einen möglichen Einfluss der farblichen Gestaltung zu untersuchen.

5.1.3.3.1 Erfassungsbogen BMS

Ein Verfahren, um den Einfluss der Arbeitsaufgabe auf die psychische Beanspruchung des menschlichen Organismus zu beurteilen, ist der Erfassungsbogen BMS [PIRi84]. Dieser verfügt über zwei anforderungsspezifische Varianten. Der BMS I wurde zur Beurteilung von Montage- und Bedienarbeitsplätzen konstruiert und der BMS II für die Beurteilung industrieller Tätigkeiten, bei denen Kontroll-, Überwachungs- und Steueraufgaben überwiegen. Beide Varianten des BMS erfassen die psychische Ermüdung (B), die Monotonie (M) und die psychische Sättigung (S). Der BMS II erfasst zusätzlich den Stress als vierte Dimension. Bis auf diesen Unterschied sind beide Fragebögen im Aufbau und in der Art der Fragenformulierung gleich. In [Heu93] wurde der BMS II für Untersuchungen an Bildschirmarbeitsplätzen eingesetzt und als mögliches Messverfahren validiert. Aus diesem Grund wurde der BMS II für die Beurteilung der Beanspruchung während der Voruntersuchung herangezogen. Er erwies sich jedoch als ungeeignet. Der Umfang wurde von den Versuchsteilnehmern als zu groß eingeschätzt und einige Items konnten nicht auf das Versuchsszenario angewendet werden. Diese Items führten zu Unsicherheit, zu Nachfragen und/oder zu ausgelassenen Antworten.

Das Ergebnis der Voruntersuchung war, dass der Erfassungsbogen BMS II eher für die Bewertung von Regelarbeitsplätzen geeignet ist. Ein Großteil der Items zielt auf Beanspruchungsfolgen ab, die bei stundenlanger bzw. täglicher Arbeit auftreten. Im Rahmen einer universitären Studie ist er daher nur schwer anwendbar.

5.1.3.3.2 Fragebogen zur Beurteilung des Trainings

Um dem eben genannten Problem entgegenzuwirken und ein situationsunabhängiges Verfahren zu erhalten, entschied sich der Versuchsleiter für einen eigenen Fragebogen (Anhang C2). Die Fragen beruhten auf den gesammelten Erfahrungen der Voruntersuchung.

Der Fragebogen besteht aus neun Items und bietet die Möglichkeit, die Beanspruchung mit einfachen Mitteln zu erfassen. Die verwendeten Items, die Zielvariablen und die jeweiligen

Antwortformate des eingesetzten Fragebogens sind der Tabelle 5-1 zu entnehmen. Die Spalte Zielvariable zeigt die gemessenen Dimensionen. Neben der Beanspruchung wird auch die Attraktivität⁵⁸ erfasst. Dieses Vorgehen wurde gewählt, um eine Korrelation zwischen der empfundenen Beanspruchung und der farblichen Gestaltung zu untersuchen.

Der Fragebogen verwendet für die Messung der Beanspruchung (Item 1–3) und für die Messung der Attraktivität (Item 4–6) jeweils drei Items. Da sich Item 4 und Item 6 auf die farbliche Gestaltung, d. h. auf die Farbe und den Kontrast, beziehen, wurden der Kontrast mithilfe des Items 5 gesondert beurteilt. Item 7 dient ebenfalls der Erfassung der Beanspruchung. Es wurde verwendet, um eine mögliche Antworttendenz der Versuchsteilnehmer auf Item 1 zu kontrollieren. Um die Motivation der Teilnehmer nicht negativ zu beeinflussen, wurde das Item bewusst nicht, wie üblich, negativ formuliert. Die Items 8 und 9 wurden verwendet, um den Versuchsteilnehmern die Möglichkeit zu geben, beide Dimensionen in Form von Stichpunkten oder Sätzen zu bewerten.

Tabelle 5-1: Fragebogen zur Beurteilung des Trainings

Nr.	Item	Zielvariable	Antwortformat
1	Wie anstrengend war das Erlernen der Software insgesamt?	Beanspruchung	sehr wenig/sehr anstrengend (15 Stufen)
2	Wie hoch waren die geistigen Anforderungen während des Erlernens?	Beanspruchung	sehr niedrig/sehr hoch (15 Stufen)
3	Wie hoch waren die körperlichen Anforderungen während des Erlernens?	Beanspruchung	sehr niedrig/sehr hoch (15 Stufen)
4	Wie angenehm ist die farbliche Gestaltung der Oberfläche insgesamt?	Attraktivität	sehr unangenehm/sehr angenehm (15 Stufen)
5	Wie angenehm ist der Kontrast der Oberfläche?	Attraktivität	sehr unangenehm/sehr angenehm (15 Stufen)
6	Wie sehr unterstützt die farbliche Gestaltung bei der Erkennung von Störungen?	Attraktivität	sehr wenig/sehr stark (15 Stufen)
7	Wie sehr mussten Sie sich anstrengen, um die Bedienung zu erlernen?	Beanspruchung	sehr wenig/sehr stark (15 Stufen)
8	Was war für Sie besonders anstrengend während der Bedienung?	Beanspruchung	Freitext
9	Was hat Sie an der Darstellung am meisten gestört?	Attraktivität	Freitext

Als Antwortformat wird eine unipolare Rating-Skala mit 15 Stufen und aufsteigender Wertigkeit verwendet. Die 15-stufige Bewertung eignet sich für eine präzise Aussage. Dabei erfolgt die Bewertung in zwei Schritten. Zuerst wird eine grobe Einschätzung getroffen (z. B. sehr niedrig–sehr hoch) und im Anschluss daran erfolgt eine dreistufige Feinjustierung (Abbildung 5-3).

⁵⁸ Der Begriff *Attraktivität* beschränkt sich in diesem Fragebogen ausschließlich auf die beiden Aspekte *Farbe* und *Kontrast*, also auf die farbliche Gestaltung. In der restlichen Arbeit beschreibt die Attraktivität eine Dimension der UX und bezieht sich auf das gesamte Produkt, d. h. auf Funktion und Design.

Ein Nachteil des Fragebogens war, dass die Beanspruchung nur global bewertet werden konnte. Die Dimensionen *Ermüdung*, *Monotonie*, *Sättigung* oder Stress konnten nicht gesondert erfasst werden – wie es beispielsweise bei dem BMS II der Fall ist. Dieser Nachteil wurde jedoch durch die Erfassung der Befindlichkeit reduziert. In weiterführenden Studien konnte gezeigt werden, dass sich Fragebögen zur Skalierung der Befindlichkeit, auch indirekt zur Beanspruchungsmessung eignen [Ape86], [Rau96]. Dieser Ansatz wurde in der vorliegenden Arbeit verfolgt.

Wie anstrengend war das Erlernen der Software insgesamt?														
sehr wenig			wenig			mittel			anstrengend			sehr anstrengend		
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Abbildung 5-3: Darstellung Antwortformat (exemplarisch Frage 1)

5.1.3.4 Fragebogen zur Beurteilung der Befindlichkeit

Die Stimmung oder Befindlichkeit kennzeichnet das nicht beobachtbare Verhalten eines Individuums, d. h. was aktuell im Inneren erlebt und subjektiv empfunden wird [SSN+94]. Wie in Kapitel 3.1.3 gezeigt wurde, kann sich beispielsweise eine länger anhaltende, niedergedrückte Stimmung negativ auf die Gesundheit auswirken.

Es existiert eine Vielzahl von Fragebögen zur Skalierung⁵⁹ der Befindlichkeit – Tabelle 5-2 zeigt hierzu eine Auswahl an Fragebögen. Neben dem Namen des Fragebogens, wird die gemessene Dimension aufgelistet und die Anzahl der Items gezeigt, welche zur Messung dieser Dimensionen verwendet werden. Die Zahl in der Klammer steht für die Gesamtanzahl der Items im Fragebogen.

⁵⁹ Skalierungsverfahren werden in der empirischen Sozialforschung überwiegend zur Messung von Einstellungen angewendet. Dabei werden zumeist mehrere Items verwendet, um verborgene und komplexe Sachverhalte aufzuzeigen und Messfehler zu reduzieren.

Tabelle 5-2: Auswahl an Fragebögen zur Skalierung der Befindlichkeit

Name	gemessene Dimensionen	Anzahl Items
<i>Positive and Negative Affectivity Scale</i> (PANAS) [WCT88], [KEK+96]	<i>positive and negative affectivity</i>	jeweils 10 (20)
<i>Swedish Occupational Fatigue Inventory</i> (SOFI) [ÅGK97], [ÅhGa98]	<i>lack of energy, physical exertion, physical discomfort, lack of motivation, sleepiness</i>	jeweils 5 (25)
Leipziger Stimmungsbogen (LSB) [HHB02]	Freude, Aktivität, Gereiztheit, Erschöpfung, Apathie	jeweils 6 (30)
Eigenzustandsskala (EZ-Skala) [Nit71]	Anstrengungsbereitschaft, Kontaktbereitschaft, Soziale Anerkennung, Selbstsicherheit, Stimmungslage, Spannungslage, Erholtheit, Schläfrigkeit	abhängig von der Dimension (40)
Mehrdimensionaler Befindlichkeitsfragebogen (MDBF) [SSN+97]	Gute-Schlechte Stimmung, Wachheit-Müdigkeit, Ruhe-Unruhe	jeweils 8 (24)

Um den passenden Fragebogen für die Bedienoberflächenstudie zu finden, wurden zwei Fragebögen ausgewählt und in der Voruntersuchung auf ihre Anwendbarkeit getestet. Die Wahl fiel auf die Eigenzustandsskala (EZ-Skala) und den Mehrdimensionalen Befindlichkeitsfragebogen (MDBF). Ein Vorteil beider Fragebögen war deren Mehrdimensionalität, d. h. neben der Stimmung wurden auch andere Aspekte (siehe Tabelle 5-2) erfasst. Des Weiteren besaß die EZ-Skala eine gute Dimensionsvariabilität und der MDBF hatte Vorteile im Bereich der Versuchsökonomie. Im Folgenden werden beide Fragebögen näher vorgestellt und die Vor- sowie die Nachteile abgewogen.

5.1.3.4.1 Die Eigenzustandsskala

Die EZ-Skala ist ein Verfahren zur Erfassung der aktuell erlebten Befindlichkeit⁶⁰ [NiUd76]. Sie stellt, neben dem eindimensionalen Ansatz, welcher z. B. der PANAS innewohnt, ein mehrdimensionales Verfahren dar [HDP+12]. Die EZ-Skala besteht aus 40 Adjektiven, zu denen jeweils eine unipolare Rating-Skala mit sechs Antwortmöglichkeiten vorgegeben wird (Abbildung 5-4).

Wie bereits erwähnt, ist ein großer Vorteil der EZ-Skala die Dimensionsvariabilität. Die 40 Adjektive (Items) können unterschiedlich verteilt werden und ergeben eine dreistufige Faktorthierarchie mit maximal 14 Faktoren. Je nach Betrachtungstiefe/-ebene, bietet die EZ-Skala die Möglichkeit, eine Aussage über zwei, vier oder acht Dimensionen der Befindlichkeit zu treffen. Als Beispiel bildet die oberste Ebene die Dimensionen *Motivation* und *Beanspruchung* ab. Da auf jeder Ebene alle 40 Adjektive aufgeteilt werden, fließen 20 Adjektive in

⁶⁰ Die Befindlichkeit wird in dem Verfahren als Eigenzustand bezeichnet.

jede Dimension ein. Auf der mittleren Ebene sind es vier Dimensionen mit jeweils zehn Adjektiven usw.⁶¹

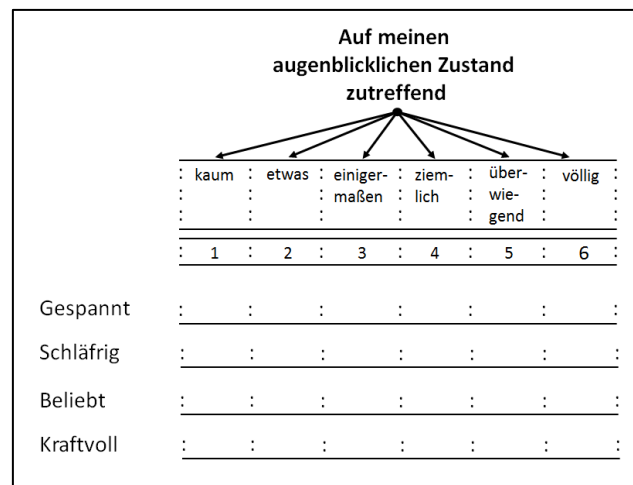


Abbildung 5-4: Auszug aus der EZ-Skala

Neben diesem Vorteil hatte die EZ-Skala auch zwei Nachteile. Wie in [ApHä84] gezeigt wird, weist sie einige Schwächen bei der internen Konsistenz der Messwerte auf. Zudem ist die Umrechnungsprozedur, die für die Auswertung notwendig ist, relativ kompliziert und aufwendig. Für die Bewertung der Daten müssen die Merkmalsausprägungen vorab standardisiert⁶², aufsummiert und anschließend auf einer gesonderten Skala abgebildet werden.⁶³ Die genannten Schwächen waren der Grund dafür, weshalb die EZ-Skala nicht für die Skalierung der Befindlichkeit in der Bedienoberflächenstudie verwendet wurde.

5.1.3.4.2 Der Mehrdimensionale Befindlichkeitsfragebogen

Der Mehrdimensionale Befindlichkeitsfragebogen (MDBF) verwendet ebenfalls eine Adjektivliste zur Erfassung der Befindlichkeit. Der Fragebogen besteht in der Langform aus 24 Items, die auf drei Dimensionen aufgeteilt werden:

- Gute-Schlechte-Stimmung (GS),
- Wachheit-Müdigkeit (WM) und
- Ruhe-Unruhe (RU).

Tabelle 5-3 zeigt die Bedeutung der einzelnen Dimensionen. Jede Dimension wird mit acht Items gemessen, wobei sich diese acht Items aus vier positiv und vier negativ orientierten Items zusammensetzen. Dadurch ist eine getrennte Untersuchung des positiven und des negativen Aspekts jeder Dimension möglich [HDP+12].

⁶¹ Für eine umfassende Beschreibung des Fragebogens und der Aufteilung der Items auf die drei Ebenen wird auf [NiUd76] verwiesen.

⁶² Normwertskala Stanine (STANDARD NINE)

⁶³ Für nähere Information zu diesem Verfahren wird auf [Nit71] verwiesen.

Tabelle 5-3: Interpretation der Dimensionen des MDBF nach [SSN+97]

Dimension	Bedeutung
GS	Ein hoher Skalenwert deutet auf eine positive Stimmungslage hin. Die Person fühlt sich wohl, ist froh und zufrieden. Niedrige Werte sprechen für eine trübsinnige und unzufriedene Person.
WM	Ein hoher Skalenwert kennzeichnet eine wache und ausgeruhte Person. Sie fühlt sich frisch und munter. Demnach fühlt sich eine Person mit niedrigem Wert eher müde, schläfrig und schlapp.
RU	Ein hoher Skalenwert entspricht einer ruhigen und gelassenen Person. Im Gegensatz dazu fühlt sich eine Person mit niedrigem Wert nervös und ist innerlich angespannt.

Wie auch bei der EZ-Skala, handelt es sich bei dem Antwortformat um eine unipolare Rating-Skala. Im Unterschied zur EZ-Skala bietet der MDBF nur fünf Antwortmöglichkeiten (Abbildung 5-5).

Im Moment fühle ich mich...	über- haupt nicht				sehr
	1	2	3	4	5
zufrieden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ausgeruht	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ruhelos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Abbildung 5-5: Auszug aus dem MDBF (Kurzform A)

Neben der Langform gibt es den MDBF auch in zwei Kurzformen mit jeweils zwölf Items (Anhang C3 und Anhang C4). Die 24 Adjektive der Langform sind auf beide Kurzformen aufgeteilt. Jede Dimension wird mit vier Items erfasst. Diese setzen sich aus zwei positiv und zwei negativ orientierten Items zusammen, wodurch der o. g. Vorteil erhalten bleibt. Durch die beiden Kurzformen wird der MDBF zu einem sehr ökonomisch einsetzbaren Instrument und eignet sich für die Prä-Posttest-Methode. Der Vorteil an dieser Methode ist, dass durch die Datenerfassung an zwei aufeinanderfolgenden Messpunkten, der Einfluss der Bedienoberfläche mittels Differenzbildung der Merkmalsausprägungen ermittelt werden kann. Der erreichbare Wert je Dimension liegt in beiden Kurzformen zwischen vier und 20, der Gesamtwert in der Langform variiert zwischen zwölf und 60.

Die Voruntersuchung hat gezeigt, dass die Ergebnisse des MDBF verlässlich sind und der MDBF für den Vergleich von Gruppen eingesetzt werden kann.⁶⁴ Des Weiteren lässt der MDBF auch Rückschlüsse auf die Beanspruchung der Versuchsteilnehmer zu. Bei dem MDBF lässt eine negative Veränderung der Dimension *WM* Rückschlüsse auf die Ermüdung zu, abfallende Skalenwerte in der Dimension *GS* sprechen für eine psychische Sättigung und sinkende Skalenwerte in der Dimension *RU* zeigen eine Stressreaktion.

⁶⁴ siehe hierzu auch [WGK+08]

Die Möglichkeit der indirekten Beanspruchungsmessung und die gute Versuchsökonomie, sprechen beide für die Verwendung des MDBF und haben letztendlich dazu geführt, dass er für die Beurteilung der Befindlichkeit in der Bedienoberflächenstudie eingesetzt wurde.

5.1.3.5 Fragebogen zur Beurteilung der Bedienungssicherheit

Um die Bedienungssicherheit der Versuchsteilnehmer zu erfassen, wurde ein weiterer Fragebogen erstellt (Anhang C5). Bei der Erstellung wurde darauf geachtet, dass die Bearbeitungszeit möglichst kurz ist, da der Fragebogen zur schnellen Erfassung gedacht war und an verschiedenen Stellen zum Einsatz kam. Aus diesem Grund beinhaltet der Fragebogen nur ein Item. Tabelle 5-4 zeigt das Item, die Zielvariable und das Antwortformat. Wie auch bei dem Fragebogen zur Beurteilung des Trainings, wird als Antwortformat eine 15-stufige unipolare Rating-Skala verwendet. Um eine Veränderung der Bedienungssicherheit über den Studienverlauf feststellen zu können, wurde die Bedienungssicherheit an vier Messpunkten (MP1–MP4) aufgezeichnet (Kapitel 5.1.5.2).

Da die Teilnehmer bei der Beantwortung nicht abschätzen konnten, inwieweit sie die notwendigen Fähigkeiten bereits erlernt hatten, wurden die Versuchsteilnehmer vor der ersten Beurteilung darauf hingewiesen, dass sich ihre Antwort nur auf die bereits erlernten und geübten Szenarien bezog. Damit wurde versucht, ausschließlich die aktuelle Bedienungssicherheit zu erfassen. Unsicherheiten in Bezug auf zukünftige Ereignisse sollten nicht berücksichtigt werden.

Tabelle 5-4: Fragebogen zur Beurteilung der Bedienungssicherheit

Nr.	Item	Zielvariable	Antwortformat
1	Wie sicher fühlen Sie sich aktuell mit der Bedienung der Simulation?	Bedienungssicherheit	sehr unsicher/sehr sicher (15 Stufen)

5.1.3.6 Handout zur Fehlerdiagnose

Während der Eingewöhnungsphase stand den Versuchsteilnehmern unterstützend ein Handout zur Fehlerdiagnose zur Verfügung. Dieses Handout fasste den Umgang mit den verschiedenen Unregelmäßigkeiten übersichtlich zusammen (Anhang B).

5.1.3.7 myGaze® Eyetracker

Die aufgezeichneten Augenbewegungen dienten der Erfassung der visuellen Wahrnehmung der Versuchsteilnehmer und wurden zusätzlich verwendet, um die vordefinierten Leistungskennzahlen (Kapitel 3.2.1 und 5.1.4) zu erfassen. Dabei wurden durch den Eyetracker die

Fixationsdauer⁶⁵, der Fixationsort und die Blickpfade aufgezeichnet. Bei dem verwendeten Eyetracker handelte es sich um ein System der Fa. myGaze®. Dieser zeichnet die Blickpfade als eine durchgehende Trajektorie⁶⁶ auf. An den Fixationsorten erzeugt die Software einen Kreis, dessen Umfang die Dauer der Fixation kennzeichnet (Abbildung 5-6). Je größer der Kreis ist, desto länger hat der Teilnehmer diese Stelle beobachtet.



Abbildung 5-6: Darstellung der Fixationsdauer, des Fixationsorts und der Blickpfade

Ausschlaggebend für die Verwendung dieses Systems waren zum einen die Robustheit und zum anderen die einfache Handhabung. Der Eyetracker funktioniert über ein Infrarotmesssystem und wird direkt am Bildschirm befestigt. Da keine Kamera auf die Augen der Versuchsteilnehmer ausgerichtet und eingestellt werden muss, wird ein aufwendiges Kalibrieren hinfällig. Des Weiteren ist das System unempfindlich gegenüber personenbedingter Besonderheiten. So war es beispielsweise kein Problem, wenn der Versuchsteilnehmer eine Brille trug. Dies war aus versuchsökonomischen Gründen sehr nützlich und sprach ebenfalls für das System.

Demgegenüber stand ein nicht unerheblicher Nachteil, welcher zu Beginn der Studie durch den Versuchsleiter nicht ausreichend bedacht wurde. Der Umfang der ausgegebenen Daten ist sehr gering und eher qualitativer Natur. Die Augenbewegungen konnten mithilfe des mitgelieferten Morae® Recorders (Version 3.3.3) nur in Form einer Videodatei erfasst werden. Aus diesem Grund mussten sie anschließend mit dem Morae® Manager (Version 3.3.3) händisch ausgewertet werden. Das Problem daran war, dass der Morae® Manager zwar die Wiedergabe der Projektdatei ermöglichte, aber nur wenige Blickdaten zur Verfügung stellte. Bei den ausgegebenen Daten handelte es sich ausschließlich um den Fixationsort und die

⁶⁵ Die Informationsaufnahme ist nur bei ruhendem Auge möglich. Diese Ruheperiode wird als Fixation bezeichnet. Sie ermöglicht erst die Fokussierung von Sehobjekten. Hierbei werden nur Ruheperioden von mindestens 100 ms gewertet. Die Augenbewegungen in Form von Blicksprüngen zwischen den Fixationen werden als Sakkaden bezeichnet. Während einer Sakkade ist die visuelle Wahrnehmung merklich herabgesetzt [KoHo14].

⁶⁶ Hierbei handelt es sich um den Bewegungspfad eines Objekts während dessen Laufzeit.

Fixationsdauer. Da die Bezugsbasis jedoch nicht bekannt war, konnten diese Koordinaten nicht ohne weiteres verwendet werden.⁶⁷ Dies erschwerte die nachträgliche Markierung von Interessengebieten (*Areas of Interest*, AOI) erheblich, wodurch Untersuchungen im Sinne des SEEV-Modells (Kapitel 2.2.1.3.2) unmöglich wurden. Zudem war eine automatisierte Auswertung der Fehlerentdeckungszeiten nicht möglich. Um die Fehlerentdeckungszeiten der Versuchsteilnehmer erfassen zu können, wurden die Augenbewegungen beobachtet und die Zeiten – mithilfe einer Stoppuhr – manuell gestoppt. Dadurch ergaben sich Unschärfen in der Auswertung der Fehlerentdeckungszeiten, auf die in Kapitel 5.2.1.5.2 noch näher eingegangen wird.

5.1.3.8 Webcam

Zusätzlich zu dem Eyetracker wurde eine Logitech C270 Webcam verwendet. Die Videoaufzeichnung wurde benutzt, um etwaige Verhaltensauffälligkeiten während des Hauptversuchs aufzuzeichnen. Hierzu zählten z. B. Unsicherheit und Überforderung der Versuchsteilnehmer. Des Weiteren wurde mithilfe der Webcam die Ablenkung durch die Nebenaufgabe dokumentiert.

5.1.3.9 Headset

Um den Versuchsteilnehmern die Möglichkeit zu geben ihre Gedanken direkt zu verbalisieren (lautes Denken) wurde ein Logitech PC 960 Headset verwendet. Die Verbalisierung mittels Headset hatte den Vorteil, dass das Problemlöseverhalten der Versuchsteilnehmer nicht beeinflusst wird [ErSi80]. Ziel dieses Vorgehens war es, ein Bild davon zu bekommen, welche kognitiven Prozesse während des Hauptversuchs in den Versuchsteilnehmern abliefen. Die Teilnehmer sollten alles ansprechen, was ihnen während der Überwachung des Bedienbereichs auffiel und durch den Kopf ging. Hierzu zählten auch Probleme mit der Bedienung und den Versuchsteilnehmern unbekannte Zustände der Simulation.

5.1.3.10 Nebenaufgabe

Nebenaufgaben (*secondary tasks*) werden in experimentellen Untersuchungen eingesetzt, um die Reservekapazitäten (*spare capacity*, [Bro65]) der Versuchsteilnehmer zu erfassen. In Kapitel 2.2.1.3.2 wurde gezeigt, dass eine Vielzahl von Ressourcen für die Aufnahme und die Verarbeitung der Informationen vorhanden sind. Diese Ressourcen können gleichzeitig genutzt werden, sind aber jede für sich begrenzt. Greifen Aufgaben auf dieselbe Ressource zu, beeinflusst die Leistung der einen Aufgabe zwangsläufig die Leistung der anderen Aufgabe.

⁶⁷ siehe hierzu auch Kapitel 5.1.3.11

Die Nebenaufgabe in der Bedienoberflächenstudie wurde auf einem Laptop angezeigt und bearbeitet. Der Laptop wurde neben dem Versuchsteilnehmer aufgestellt, so dass er bequem während des Simulationslaufs bedient werden konnte (Abbildung 5-7). Die Nebenaufgabe bestand aus dem Surrogate Reference Task (SURT). Der SURT ist eine visuelle Suchaufgabe, die optische und kognitive Ablenkung hervorrufen kann. Hierdurch kann die Reservekapazität der Versuchsteilnehmer erfasst werden [WyRi08].

Die Teilnehmer hatten die Aufgabe, einen Zielreiz (Target) unter mehreren Ablenkern (Distraktoren) zu finden. Der Bildschirm des Laptops war in zwei Bereiche unterteilt (Abbildung 5-7) und die Auswahl der Position des Zielreizes erfolgte mit den Pfeiltasten. Die Schwierigkeit der Aufgabe konnte durch die Größe der Kreise variiert werden. In der Studie wurde für den Zielreiz ein Durchmesser von 15 mm und für die Ablenker ein Durchmesser von 7.5 mm gewählt. Diese Größen wurden durch den Versuchsleiter als ausreichend definiert. Die Anzahl der Störreize pro Bildschirm wurde auf $N = 50$ festgelegt.

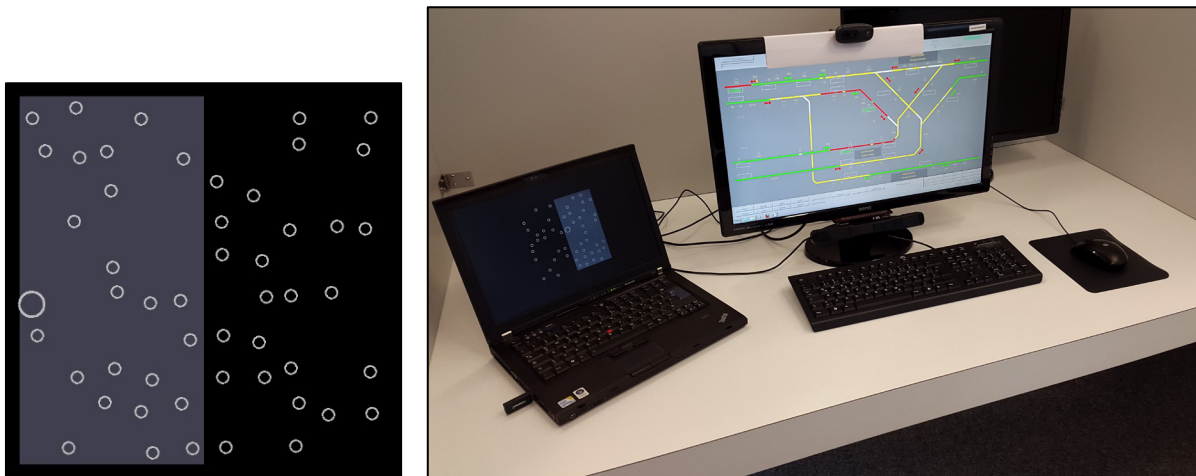


Abbildung 5-7: Darstellung der Nebenaufgabe auf dem Laptop-Bildschirm (links) und Position der Nebenaufgabe während des Versuchs (rechts)

5.1.3.11 Tool zur Erstellung von Heatmaps

Wie bereits erwähnt, konnten die Koordinaten der Fixationsorte nicht direkt verwendet werden. Die Koordinaten der einzelnen Fixationsorte werden durch die Recorder-Software in einem kartesischen Koordinatensystem abgebildet – vermutlich 1. Quadrant. Das Problem daran war, dass die Definition des Nullpunktes und der x- sowie der y-Achse programmintern geschahen. Dadurch konnte kein externer Koordinatenbezug der Fixationspunkte hergestellt werden. Um dennoch mögliche Auffälligkeiten im Gesamtblickverhalten der Versuchsteilnehmer zu erkennen, wurden Heatmaps aus den Blickdaten erstellt. Dabei wurde ein in Java programmiertes Tool verwendet. Das Tool wurde von der Fa. Siemens entwickelt und arbeitet browserbasiert (Anhang E).

5.1.3.12 Visualisierung der Mausbewegungen mit IOGraph

Für die Visualisierung der Mausbewegungen wurde ein einfaches Mouse-Tracking-Tool verwendet. Bei diesem Tool handelte es sich um IOGraph der Fa. IOGraphica [ZeSh16]. Das Tool zeichnet eine durchgehende Trajektorie auf und erzeugt an dem Verweilort des Mauszeigers zwei konzentrische Kreise, die die Verweildauer des Mauszeigers darstellen. Bei dem inneren, ausgefüllten Kreis kennzeichnet die Fläche die Verweildauer und bei dem äußeren Kreis ist der Durchmesser proportional zur Verweildauer [Pac15].

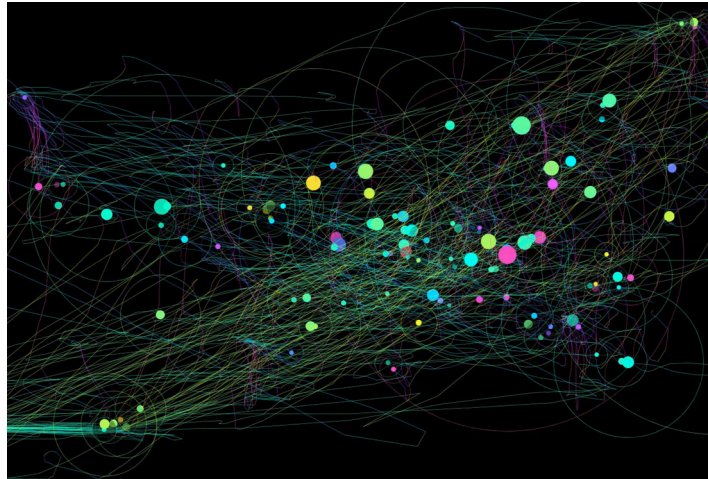


Abbildung 5-8: Ausschnitt aus der JPEG-Datei, die durch das Tool *IOGraph* erzeugt wird

Das Tool gibt keine Koordinaten oder sonstige Daten aus. Nach Beendigung der Aufzeichnung wird automatisch eine JPEG-Datei erstellt (Abbildung 5-8). Demnach ist das Verfahren nur für eine qualitative Beurteilung geeignet.

5.1.4 Operationalisierung der Bedienereffizienz

Im Rahmen der Bedienoberflächenstudie wurden verschiedene vordefinierte Leistungskennzahlen erfasst (Tabelle 3-1). Diese Leistungskennzahlen spannen ein Modell auf, welches die Faktoren *Leistungsfähigkeit* und *Handlungssicherheit* betrachtet. Diese beiden Faktoren wurden in der vorliegenden Arbeit dazu verwendet, die Bedienereffizienz im Stellwerk versuchsweise zu operationalisieren. Nachfolgend wird der theoretische Ansatz vorgestellt und die beteiligten Faktoren erklärt.

5.1.4.1 Theoretischer Ansatz

Wie in Kapitel 3.2 gezeigt wurde, wird die Bedienereffizienz als Fähigkeit beschrieben, Bedienhandlungen räumlich und zeitlich so zu bündeln, dass sich ein zeitsparender Bedienungsablauf ergibt. Dabei kann die Bedienereffizienz zum einen durch das Bedienoberflächendesign und zum anderen durch das Situationsbewusstsein des Bedieners beeinflusst werden [Pac15]. Diese beiden Einflüsse werden in der Bedienoberflächenstudie über die

Aspekte *Leistungsfähigkeit* und *Handlungssicherheit* beschrieben. Abbildung 5-9 zeigt das Modell und gibt eine Übersicht über die Zuordnung der Leistungskennzahlen sowie die zusätzlich wirkenden Faktoren.

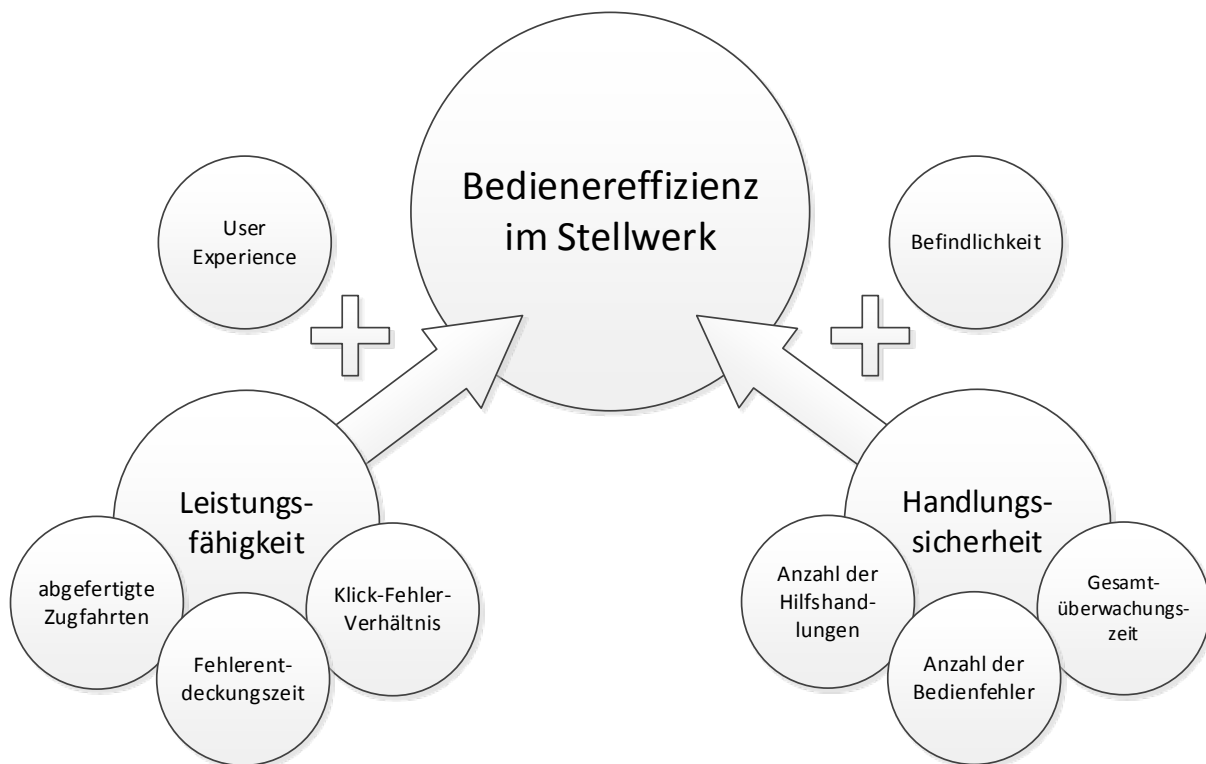


Abbildung 5-9: Konzept zur Operationalisierung der Bedienereffizienz

In Kapitel 2 der Arbeit wurde gezeigt, dass die farbliche Gestaltung einen positiven Einfluss auf den Bediener haben kann. Die als positiv empfundene Ästhetik kann beispielsweise die subjektive Wahrnehmung der Usability erhöhen [TKI00]. Dadurch wird die Attraktivität des Systems verbessert und die Akzeptanz des Bedieners gegenüber dem System steigt [Mos12]. Das wiederum, kann sich in einer besseren Leistungsbereitschaft niederschlagen [MoTh10]. Um diesen Zusammenhang genauer untersuchen zu können, wurden drei Leistungskennzahlen definiert. Diese wurden verwendet, um den Faktor *Leistungsfähigkeit* zu operationalisieren. Hierzu zählten:

- die Anzahl der abgefertigten Zugfahrten (Kapitel 5.1.4.2.1),
- die Fehlerentdeckungszeit (Kapitel 5.1.4.2.2) und
- das Verhältnis zwischen der Klickzahl und den gemachten Fehlern in der Nebenaufgabe (Kapitel 5.1.4.2.3).

Hierbei handelte es sich ausschließlich um objektive Leistungsmaße, mit denen eine vergleichbare Basis geschaffen werden sollte. Anschließend wurden die drei Leistungskennzahlen mit der UX in Bezug gesetzt, um den o. g. Zusammenhang zu überprüfen.

Wie bereits erwähnt, ist der zweite Faktor, welcher die Bedienereffizienz beeinflusst, das Situationsbewusstsein. In Kapitel 2.1.1.2.1 wurde gezeigt, dass ein angemessenes Situationsbewusstsein immer dann vorliegt, wenn der Mensch sich bei der Aufgabenerfüllung zu jedem Zeitpunkt darüber im Klaren ist, was gerade passiert, warum es passiert und wie sich die Situation weiter entwickeln wird [Man08]. Mangelndes Situationsbewusstsein zeigt sich vor allem dann, wenn die Sicherungstechnik ausfällt oder unerwartet nicht einwandfrei funktioniert.

Demnach muss für den Aufbau eines angemessenen Situationsbewusstseins auch ein ausreichendes Systemverständnis vorhanden sein und der Bediener muss sich seiner Aufgabe und der notwendigen Bedienhandlungen zu jeder Zeit sicher sein. Das setzt voraus, dass der Bediener die Bedienung der Anlage beherrscht. Dieser Umstand wird in der vorliegenden Arbeit als Handlungssicherheit definiert. *NEBYLIZYN* versteht unter der Handlungssicherheit die Fähigkeit, trotz Situationskomplizierung, innerhalb vorgegebener Zeiträume beanspruchungsoptimal und anforderungsgerecht zu reagieren [Neb63]. Um den Faktor *Handlungssicherheit* zu operationalisieren, wurden drei weitere Leistungskennzahlen definiert. Hierzu zählten:

- die Anzahl der Hilfshandlungen (Kapitel 5.1.4.3.1),
- die Anzahl der Bedienfehler (Kapitel 5.1.4.3.2) und
- die Gesamtüberwachungszeit (Kapitel 5.1.4.3.3).

In [Rau96] wird kritisiert, dass die Bewertung der Handlungssicherheit oft unter unzulässiger Reduktion auf die anforderungsgerechte Ausführung der Tätigkeit erfolgt. Dabei wird größtenteils ergebnisorientiert vorgegangen, d. h. Kriterien für Handlungsunsicherheit sind gemachte Fehler und/oder nicht erbrachte Leistungen. Der Aspekt der beanspruchungsoptimalen Ausführung bleibt dadurch außen vor. Um diesem Problem entgegenzuwirken, wurden neben Art und Anzahl von Bedienfehlern auch korrekte Hilfshandlungen gezählt und die Gesamtüberwachungszeit der Versuchsteilnehmer auf dem Bildschirm erfasst. Dabei ermöglichte die erfasste Anzahl der Hilfshandlungen eine Aussage über das Systemverständnis der Teilnehmer und die Gesamtüberwachungszeit über die Verteilung der Ressourcen auf die Haupt- und Nebenaufgabe.

Um der Definition nach [Neb63] vollständig gerecht zu werden, wurden die Variablen mit den Ergebnissen der erfassten Befindlichkeit in Zusammenhang gebracht. Dadurch wurde der Aspekt der Beanspruchung in die Betrachtung integriert. Die Überlegung war, dass Versuchsteilnehmer, die die ihnen gestellte Aufgabe gut erfüllten, nur als handlungssicher galten, wenn ebenfalls eine niedrige Abweichung der Belastung vom Optimum vorlag. Dabei wurden bewusst die Ergebnisse des MDBF und nicht die des Fragebogens zur Beurteilung des Trainings herangezogen, da der MDBF auch indirekte Rückschlüsse auf die verschiedenen Dimensionen der Beanspruchung zuließ (Kapitel 5.1.3.4.2).

5.1.4.2 Erfassung der Leistungsfähigkeit

Im Folgenden werden die drei Leistungskennzahlen erläutert, mit deren Hilfe die Leistungsfähigkeit der Versuchsteilnehmer beurteilt wurde.

5.1.4.2.1 Anzahl der abgefertigten Zugfahrten

Die Anzahl der abgefertigten Zugfahrten wurde erfasst, um die Leistung der Versuchsteilnehmer objektiv beurteilen zu können. Es wurde vermutet, dass Versuchsteilnehmer, die ein ausreichendes Systemverständnis haben und in der Bedienung gut geübt sind, auch einen hohen Durchsatz haben. Demnach ist die Anzahl stark von der Handlungssicherheit der Versuchsteilnehmer abhängig. Diese Korrelation wird im Rahmen der Studie überprüft.

Um die abgefertigten Zugfahrten erfassen zu können, wurden vorab die Bedingungen für eine Zählung definiert. Das verwendete Lupenbild besitzt drei „Einbruchsmöglichkeiten“ für Zugfahrten und vier mögliche Fahrstraßen, auf denen die Züge den Bedienbereich wieder verlassen können. Die Fahrstraße über das Signal S03 wurde in der Simulation nicht planmäßig befahren, weshalb nur die übrigen Signale S06, S20 und S29 als Austrittssignale festgelegt wurden. Die Bedingungen für eine Zählung waren:

- das komplette Passieren des Signals durch den Zug und
- der Signalhaltfall des Austrittssignals.

5.1.4.2.2 Fehlerentdeckungszeit

Den Bedienbereich zu überwachen und Unregelmäßigkeiten zu entdecken, war die Hauptaufgabe der Versuchsteilnehmer. Um eine Aussage über die Reaktionsgeschwindigkeit der Versuchsteilnehmer treffen zu können und mögliche Unterschiede bei den verwendeten Bedienoberflächen feststellen zu können, wurde die Zeit vom Eintreten bis zum Erkennen einer Unregelmäßigkeit gemessen. Diese Zeit wurde als Fehlerentdeckungszeit bezeichnet.

5.1.4.2.3 Verhältnis zwischen Klickzahl und gemachten Fehlern in der Nebenaufgabe

Bei dem dritten Leistungsmaß handelte es sich um das Verhältnis zwischen der Klickzahl und den gemachten Fehlern in der Nebenaufgabe. Die Variable wurde als Klick-Fehler-Verhältnis (KfV) bezeichnet. Das KfV errechnet sich mit der Formel 5-1 aus Abbildung 5-10.

$$KfV = \frac{\text{Klicks}}{\text{Fehler}}$$

Abbildung 5-10: Bestimmung des Klick-Fehler-Verhältnisses (Formel 5-1)

Je größer das Verhältnis ist, desto besser war das Ergebnis der Versuchsteilnehmer in der Nebenaufgabe.

5.1.4.3 Erfassung der Handlungssicherheit

Im Folgenden werden die drei Leistungskennzahlen erläutert, mit deren Hilfe die Handlungssicherheit der Versuchsteilnehmer erfasst wurde.

5.1.4.3.1 Anzahl der Hilfshandlungen

Die Anzahl der korrekten Hilfshandlung wurde erfasst, um eine Vorstellung darüber zu bekommen, wie gut die Versuchsteilnehmer die Bedienung der Simulation beherrschten. Da korrekt und gezielt ausgeführte Hilfshandlungen ein Zeichen für ein ausreichendes Systemverständnis waren, wurden sie nicht zur Beurteilung der Leistungsfähigkeit verwendet, sondern bildeten einen essentiellen Bestandteil bei der Beurteilung der Handlungssicherheit.

Wie bereits erwähnt, wurden nur korrekte Hilfshandlungen gezählt. Falsche oder unnötige Hilfshandlungen zählten zu den Bedienfehlern (Kapitel 5.2.1.5.5). Da die Anzahl der notwendigen Hilfshandlungen für die verschiedenen Unregelmäßigkeiten stark variierte und vom Trainingsstand der Versuchsteilnehmer abhängig war, wurden nur Hilfshandlungen gezählt, durch die eine Zugfahrt möglich wurde. Das heißt, dass nur das erfolgreiche Stellen eines Vorsichtssignals als Hilfshandlung gezählt wurde.

Da die Unregelmäßigkeiten durch den Versuchsleiter vorgegeben wurden (Anhang D), kamen sechs Signale für die Zählung in Betracht. Tabelle 5-5 zeigt die Ergebnisse der Zählung an einem Beispiel – hier exemplarisch an dem Versuchsteilnehmer 49. Zu sehen sind die sechs Signale und der Ursprung der VS-Fahrt. Um den Ursprung des Vorsichtssignals besser nachvollziehen zu können, wurden die VS-Fahrten in *VS-Fahrt Weiche* und *VS-Fahrt Signal* unterteilt. Der erste Wert zeigt die gezählten Vorsichtssignale aufgrund einer Gleisfreimelde-störung und der letzte Wert die gezählten Vorsichtssignale aufgrund einer Signalstörung.

Tabelle 5-5: Übersicht über die gezählten Hilfshandlungen (exemplarisch Versuchsteilnehmer 49)

	S01	S11	S22	S24	Summe
VS-Fahrt Weiche	1	2	1	2	6
	S01	S23	S24	S25	Summe
VS-Fahrt Signal	1	2	2	1	6

5.1.4.3.2 Anzahl der Bedienfehler

Nach [Rau96] sind Art und Anzahl von Fehlern während der Aufgabenerfüllung klassische Kriterien für die Erfassung der Handlungssicherheit. Aus diesem Grund wurden die Bedienfehler der Versuchsteilnehmer erfasst.

Zu den Bedienfehlern zählten alle falschen und unnötigen Bedienhandlungen. Um die Fülle an Bedienfehlern besser vergleichen zu können, wurden sie zu Beginn der Auswertung klassifiziert. Tabelle 5-6 zeigt die Klassifizierung der Bedienfehler.

Tabelle 5-6: Übersicht über die Klassifizierung der Bedienfehler

Bedienfehler	Erklärung
falsches VS	Teilnehmer versucht ein Vorsichtssignal einzustellen, obwohl keine Störung vorliegt und keine Fahrstraße (FS) eingestellt ist
überflüssiges VS	...versucht ein Vorsichtssignal einzustellen, obwohl keine Störung mehr vorliegt und die FS bereits eingelaufen ist
VS ohne FS	...versucht ein Vorsichtssignal einzustellen, weil eine Störung vorliegt , ohne vorher eine FS einzustellen
VS ohne Auflösung	...stellt Vorsichtssignal ohne die Fahrwegelemente vorher aufzulösen
VS statt SES	...stellt ein weiteres Vorsichtssignal, obwohl die Signalstörung bereits behoben wurde und das Signal entsperrt werden muss
WHU ohne WHA	...versucht eine gestörte Weiche umzustellen, vergisst aber die Weiche vorher aufzulösen
unnötiges WHU	...versucht eine nicht gestörte Weiche mit einer Hilfshandlung umzustellen
FS ohne WHU	...versucht eine FS einzustellen, obwohl er die Weiche noch nicht in die richtige Lage gebracht hat
unnötiges Auflösen Fahrweg-element	...versucht ein nicht verschlossenes Element aufzulösen
falsches Auflösen Fahrweg-element	...versucht ein Element fälschlicherweise aufzulösen (Element im D-Weg, aktuell eine Zugfahrt, Flankenschutz oder bereits (neue) FS eingestellt)
verursacht Fehllenkung	...lenkt Zugfahrt falsch
übersieht KF1 oder KF2	...ignoriert oder übersieht die KF-Bedienung

5.1.4.3.3 Gesamtüberwachungszeit

Die Gesamtüberwachungszeit wurde erfasst, um zu untersuchen, ob ein Zusammenhang zwischen dem Überwachungsverhalten und den Fehlerentdeckungszeiten bestand. Sie beschreibt die Gesamtblickverweildauer der Versuchsteilnehmer auf dem Bildschirm während des Hauptversuchs. Je kleiner die Gesamtüberwachungszeit war, desto weniger Zeit verbrachten die Versuchsteilnehmer mit der Überwachung des Bedienbereichs. Sie wurde durch das Aufsummieren der einzelnen Fixationszeiten (Kapitel 5.1.3.7) ermittelt.

5.1.5 Durchführung (Versuchsablauf)

Der Versuch gliederte sich in drei Teile und die Gesamtversuchsdauer betrug in etwa drei Stunden. In der Regel wurden die Daten von zwei Versuchsteilnehmern gleichzeitig erfasst. Dabei wurde ein Teilnehmer durch den Versuchsleiter betreut und ein zweiter durch eine studentische Hilfskraft. Tabelle 5-7 zeigt den Versuchsablauf mit den dazugehörigen Zeitwerten der einzelnen Phasen. Um die durchgeführte Bedienoberflächenstudie besser nachvollziehen zu können, wird im Folgenden der Inhalt und der Ablauf der einzelnen Phasen genauer beschrieben.

Tabelle 5-7: Darstellung des Versuchsablaufs

Aktion	Zeit (min)
Teil 1 (Einführungsphase)	
Begrüßung und Ausgabe der Einverständniserklärung	5
Einführungspräsentation	15
Ausgabe des Handouts (inkl. Einarbeitungszeit)	10
	30
Teil 2 (Eingewöhnungsphase mit standardisiertem Testlauf)	
MDBF Kurzform A	5
3 Trainingsblöcke je 20 min für das Erlernen der Simulation	60
⇒ Aufzeichnung Mausbewegungen in jedem Trainingsblock ⇒ vor jedem Trainingsblock Fragebogen zur Beurteilung der Bedienungssicherheit (MP1+MP2)	
Standardisiertes Szenario (Testlauf) für die Überprüfung des Wissensstandes (<i>Simulation +5 min</i>)	15
⇒ Aufzeichnung Mausbewegungen ⇒ vor dem Testlauf Fragebogen zur Beurteilung der Bedienungssicherheit (MP3)	
Fragebogenblock inkl. Pause: Beurteilung des Trainings, MDBF Kurzform B, UEQ	15
	95
Teil 3 (Hauptversuchsphase mit Interview)	
Hauptversuch (<i>Simulation +10 min</i>)	30
⇒ Aufzeichnungen Mausbewegungen ⇒ vor dem Hauptversuch Fragebogen zur Beurteilung der Bedienungssicherheit (MP4) ⇒ Aufzeichnung Augenbewegungen ⇒ Nebenaufgabe ⇒ Videoaufzeichnung ⇒ Tonaufnahme	
Interview (Befragung zum Designbedienplatz)	10
	40
Gesamtversuchsdauer	165

5.1.5.1 Einführungsphase (Dauer ca. 30 min)

Zu Beginn der Studie wurden die Versuchsteilnehmer durch den Versuchsleiter begrüßt. Danach bekamen sie eine kurze Einführung in Form einer Präsentation (Anhang A). Diese beinhaltete, neben ein paar allgemeinen Informationen zum Eisenbahnwesen und zur Stellwerkstechnik, die Vorstellung der verwendeten Simulation sowie eine erste Erklärung der ver-

schiedenen Unregelmäßigkeiten⁶⁸. Diese wurden dann an jeweils einem Beispiel erklärt. Danach bekamen die Versuchsteilnehmer das Handout zur Fehlerdiagnose (Kapitel 5.1.3.6) ausgehändigt und es folgte die Eingewöhnungsphase, welche mit einem standardisierten Testlauf endete.

5.1.5.2 Eingewöhnungsphase (Dauer ca. 90 min)

Nach einer kurzen Vorbereitungszeit, die dem Kennenlernen des Handouts diente, füllten die Teilnehmer den ersten Fragebogen aus – den MDBF in der Kurzform A (Kapitel 5.1.3.4.2). Ziel war es, den Ausgangswert der Befindlichkeit zu erfassen. Nachdem die Teilnehmer den MDBF beantwortet hatten, begann die eigentliche Eingewöhnungsphase. Hier wurde den Teilnehmern der Umgang mit der Simulation näher gebracht. Zudem sollten mögliche Unsicherheiten (z. B. Testangst), verursacht durch die apparativen Gegebenheiten und durch eine anfängliche Aufregung, eliminiert werden. Diese Unsicherheit hätte eine Veränderung der Messdaten zur Folge haben können. Aus diesem Grund wurde versucht, diese durch eine Face-to-Face-Betreuung der Versuchsteilnehmer abzubauen.

Die Eingewöhnungsphase unterteilte sich in drei Trainingsblöcke, wobei jeder Block eine Dauer von 20 min hatte. Vor dem zweiten und dritten Trainingsblock wurde der Fragebogen zur Beurteilung der Bedienungssicherheit ausgegeben und durch die Teilnehmer ausgefüllt (MP1, MP2). Diese Abfrage erfolgte, um eine Veränderung der Bedienungssicherheit über den Studienverlauf feststellen zu können. Eine Besonderheit der einzelnen Trainingsblöcke war, dass der Automatisierungsgrad stetig zunahm. Anfangs war der Zuglenkbetrieb an allen Signalen ausgeschaltet und die Versuchsteilnehmer mussten sämtliche Fahrstraßen händisch einstellen. Mit zunehmendem Fortschritt des Trainings wurde der Zuglenkbetrieb an immer mehr Signalen wieder eingeschaltet. Im dritten Trainingsblock war der Betrieb dann komplett automatisiert. Dieses Vorgehen hatte zwei Gründe. Zum einen war das Erlernen der grundsätzlichen Bedienung notwendig, weil für das Behandeln der Gleisfreimelde störung die Fahrstraßen händisch eingestellt werden mussten und zum anderen sollte ein ausreichendes Systemverständnis aufgebaut werden. Dies wurde zusätzlich durch den abnehmenden Schwierigkeitsgrad der Unregelmäßigkeiten verstärkt.⁶⁹ Die Versuchsteilnehmer wurden durch den niedrigen Automatisierungsgrad und die hohe Schwierigkeit der Unregelmäßigkeiten anfänglich stärker gefordert, bekamen am Ende jedoch mehr Raum, um das Erlernete zu vertiefen und den Lernprozess zu unterstützen.

Im Anschluss an den dritten Trainingsblock wurde das Handout zur Fehlerdiagnose wieder eingesammelt und es folgte der standardisierte Testlauf. Die Aufgabe der Versuchsteilnehmer während des Tests bestand darin, den Bedienbereich zu überwachen und auf mögliche

⁶⁸ siehe hierzu Kapitel 4.3.2

⁶⁹ Die drei Unregelmäßigkeiten wurden auf die drei Trainingsblöcke aufgeteilt. Im ersten Trainingsblock mussten die Versuchsteilnehmer den Umgang mit der Gleisfreimelde störung erlernen (Kapitel 4.3.2.3), im zweiten Trainingsblock wurde der Umgang mit der Signalstörung vermittelt (Kapitel 4.3.2.2) und im letzten Block wurde der Ausfall des ZLB simuliert (Kapitel 4.3.2.1).

Unregelmäßigkeiten zu reagieren. Er wurde unter „realen“ Bedingungen durchgeführt, d. h. es wurde keine Hilfestellung mehr geleistet. Mit dem Testlauf wurde der Trainingsstand der Teilnehmer überprüft. Zudem wurde kontrolliert, ob die Versuchsteilnehmer die notwendigen Bedienungen erlernt hatten und den Umgang mit den drei Unregelmäßigkeiten beherrschten. Die Verteilung der Unregelmäßigkeiten wird im (Anhang D) dargestellt.⁷⁰ Bevor der Testlauf gestartet wurde, wurde der Fragebogen zur Beurteilung der Bedienungssicherheit ein drittes Mal ausgefüllt (MP3).

Nach dem standardisierten Testlauf folgten ein Fragebogenblock und eine zehnminütige Pause. Der Fragebogenblock bestand aus dem Fragebogen zur Beurteilung des Trainings (Kapitel 5.1.3.3.2), dem MDBF in der Kurzform B (Kapitel 5.1.3.4.2) und dem UEQ (Kapitel 5.1.3.2). Im Anschluss daran folgte die Hauptversuchsphase.

5.1.5.3 Hauptversuchsphase (Dauer ca. 40 min)

Zu Beginn der Hauptversuchsphase wurde zum vierten und letzten Mal der Fragebogen zur Beurteilung der Bedienungssicherheit ausgefüllt (MP4). Im Abschluss daran wurden die Versuchsteilnehmer für den Hauptversuch vorbereitet. Hierfür wurde ihnen das Headset aufgesetzt, der Eyetracker eingerichtet, die Nebenaufgabe erklärt⁷¹ und die Webcam eingeschaltet. Die Versuchsdauer in der Hauptversuchsphase betrug 30 min. Die Dauer wurde bewusst so gewählt, da festgestellt wurde, dass der Konzentrationsverlust in der ersten halben Stunde am größten ist [Kra85]. Es wurde vermutet, dass in dieser Zeit am ehesten eine Veränderung der abhängigen Variablen erkennbar ist und eine unterschiedliche Wahrnehmung der beiden Bedienplätze messbar wird. Die Aufgabe der Versuchsteilnehmer lag wie bei dem standardisierten Testlauf darin, den Bedienbereich zu Überwachen und auf mögliche Unregelmäßigkeiten zu reagieren.

Im Hauptversuch gab es 28 Unregelmäßigkeiten. Diese setzten sich aus 21 ZLB-Ausfällen, vier Signalstörungen und drei Gleisfreimeldestörungen zusammen. Die Schwierigkeitsverteilung der einzelnen Unregelmäßigkeiten folgte einer „Glockenkurve“ (Anhang D). Als Beispiel trat die Gleisfreimeldestörung hauptsächlich in der Mitte des Hauptversuchs auf, da sie am schwersten zu erkennen war und von der Anzahl der Hilfshandlungen am umfangreichsten war. Dieses Vorgehen wurde bewusst gewählt, um die Beanspruchung der Versuchsteilnehmer in dieser Phase zu erhöhen. Es wurde vermutet, dass hierdurch ein Defizit im Trainingsstand zu erkennen war, was wiederum indirekte Rückschlüsse auf die Leistungsfähigkeit und die Handlungssicherheit zuließ.

⁷⁰ Die Simulation der Stadtbahn wurde vor dem standardisierten Testlauf und vor dem Hauptversuch neugestartet. Dieses Vorgehen wurde gewählt, um für alle Versuchsteilnehmer die gleichen Ausgangsbedingungen zu schaffen. Da es eine gewisse Zeit in Anspruch nahm, bis sich wieder ausreichend Züge im Netz der Stadtbahn befanden, wurde die Betriebssituation in der Simulation um eine definierte Zeit „vorgespult“. Die Zeiten sind in der Tabelle 5-7 kursiv dargestellt.

⁷¹ Die Versuchsteilnehmer konnten an dieser Stelle den Umgang mit der Nebenaufgabe testen.

Zudem wurde der Umgang mit der Nebenaufgabe beobachtet, um eine Aussage über die Ressourcenverteilung der Teilnehmer treffen zu können. Durch die Steigerung des Schwierigkeitsgrades in der Hauptaufgabe, sollte ein Leistungseinbruch in der Nebenaufgabe entstehen.

Auf eine Besonderheit bei der Abfolge der Unregelmäßigkeiten im Hauptversuch sei an dieser Stelle noch hingewiesen. Um Defizite im Überwachungsverhalten der Versuchsteilnehmer, bezogen auf bestimmte Bereiche der Bedienoberfläche, feststellen zu können, fiel der ZLB des Signals S08 in der 26 min. einmal aus. Das Signal war während des kompletten Versuchs unauffällig und lag inmitten einer Reihe anderer Elemente, von denen der Teilnehmer mehr relevante Informationen erhielt und sie deshalb auffälliger waren.

Nach dem Hauptversuch wurde ein nicht standardisiertes Interview durchgeführt. Den Versuchsteilnehmern wurde der Designbedienplatz (Kapitel 4.2.2) gezeigt. Es wurden hauptsächlich Fragen zum Kontrast und zur Fehlerwahrnehmung sowie zu störenden Eigenschaften gestellt. Die Antworten bezogen sich auf den Direktvergleich mit dem kennengelernten Bedienplatz.

5.1.6 Statistische Auswertung

Die statistische Auswertung der Daten erfolgte mit der Statistiksoftware SPSS (Version 23.0). Für alle Tests wurde das Signifikanzniveau α auf .05 gesetzt. Da Zweifel an der Gleichheit der Abstände zwischen den einzelnen Stufen der Ratingskalen bestanden, waren die Merkmale der verwendeten Ratingskalen strenggenommen ordinalskaliert. Demnach hätten die Ergebnisse mit einem Verfahren für Ordinaldaten ausgewertet werden müssen. Da dies die Verwendung von sinnvoll interpretierbaren Mittelwerten ausschließt und es sich in der Forschungspraxis eingebürgert hat, auch solche Ratingskalen⁷² als intervallskaliert aufzufassen [BoDö15], wurde dieser Ansatz in der vorliegenden Arbeit ebenfalls verfolgt. Dieses Vorgehen wurde dadurch abgesichert, dass die erhobenen Daten im strittigen Fall sowohl mit einem Verfahren für intervallskalierte Daten (parametrisches Verfahren) als auch mit einem Verfahren für ordinalskalierte Daten (nicht parametrisches Verfahren) ausgewertet wurden. Als parametrisches Verfahren kam der t-Test für unabhängige Stichproben zum Einsatz und als nicht parametrisches Verfahren der Mann-Whitney-U-Test.

Wie der Name bereits zeigt, gibt es bei dem parametrischen Verfahren gewisse Voraussetzungen (Parameter), die erfüllt sein müssen, um das Verfahren anwenden zu können. Für die Verwendung des t-Tests müssen die Daten intervallskaliert und normalverteilt⁷³ sein und die Varianzen müssen gleich sein [BoDö15]. Die Verteilungsanalyse erfolgte mit einem inferenz-

⁷² wenn sie mindestens fünf Stufen haben

⁷³ Ab einer Stichprobengröße von $N > 30$ greift das zentrale Grenzwerttheorem. Es kann davon ausgegangen werden, dass die Mittelwerte der Merkmalausprägungen unabhängig sind und annähernd einer Normalverteilung folgen.

statistischen Standardtest. Als Standardtest wurde der Shapiro-Wilk-Test verwendet, da es sich dabei um einen Test mit hoher Güte, d. h. geringem β -Fehler handelt [BoSc10]. Im Anschluss daran wurde die Varianzhomogenität mithilfe des Levene-Tests überprüft. Sofern nur die Bedingung der Normalverteilung der Merkmalsausprägungen nicht erfüllt wurde, wurden die Hypothesen trotzdem mit dem t-Test überprüft. Dieses Vorgehen wurde gewählt, da das Verfahren robust gegenüber Verletzungen der Normalverteilungsbedingung ist [BoSc10]. Voraussetzung hierfür ist, dass eine homogene Gruppengröße vorliegt – was in der Bedienoberflächenstudie der Fall war. Der Vollständigkeit halber wird erwähnt, dass für die Anwendung des Mann-Whitney-U-Tests keine besonderen Voraussetzungen erfüllt sein müssen. Zudem besitzt er bei Verletzung der Voraussetzungen für ein parametrisches Verfahren die bessere Teststärke.

Weiter wurden für die Diskussion der Ergebnisse (Kapitel 5.3) mögliche Zusammenhänge zwischen den abhängigen Variablen untersucht. Als Maß des Zusammenhangs wurden der Korrelationskoeffizient nach Pearson (r) und der Rangkorrelationskoeffizient nach Spearman (r_s) verwendet. Die Größe der Korrelation wird als Effektstärke bezeichnet. COHEN definiert für beide Koeffizienten einen möglichen Effekt als gering bzw. schwach, wenn $r/r_s \geq .10$ ist, als mittel bzw. moderat bei $r/r_s \geq .30$ und als groß bzw. stark, wenn der Effekt größer $r/r_s = .50$ ist ([Coh88], S. 82). Wie auch bei den anderen Daten, wurden jeweils ein parametrisches und ein nicht parametrisches Verfahren für die Untersuchung verwendet. Der Korrelationskoeffizient gehört zu den parametrischen Verfahren und der Rangkorrelationskoeffizient nach Spearman zu den nicht parametrischen Verfahren [BoDö15]. Dabei besitzt letzterer einen entscheidenden Vorteil. Der Rangkorrelationskoeffizient ist robust gegenüber Ausreißern⁷⁴ [NN16a]. Ausreißer können zu einer Invertierung des Zusammenhangs führen, was zu einer Fehlinterpretation führt. Aus diesem Grund wurde das Ergebnis des Rangkorrelationskoeffizienten in der vorliegenden Arbeit stärker gewichtet. Tabelle 5-8 fasst die verwendeten Tests zusammen, gibt einen Überblick über die notwendigen Voraussetzungen für deren Anwendung und zeigt die Besonderheiten des Verfahrens bzw. des Koeffizienten.

Tabelle 5-8: Gegenüberstellung der verwendeten statistischen Verfahren

Kategorie	parametrisch	nicht parametrisch
Verfahren	t-Test	Mann-Whitney-U-Test
Korrelationskoeffizient	Pearson (r)	Spearman (r_s)
Voraussetzung	<ul style="list-style-type: none"> • intervallskaliert • normalverteilt (Levene-Test) • homogene Varianzen (Shapiro-Wilk-Test) 	keine
Besonderheit	<ul style="list-style-type: none"> • robust gegenüber Verletzung der Normalverteilungsbedingung • Ausreißer können zur Invertierung der Korrelationsrichtung führen 	<ul style="list-style-type: none"> • bessere Teststärke, wenn Voraussetzungen für parametrisches Verfahren nicht erfüllt • robust gegenüber Ausreißern

⁷⁴ Beobachtungswert, der scheinbar nicht zu den übrigen Werten passt. Ausreißer wirken sich sehr stark auf das arithmetische Mittel und auf Streuungsmaße wie die Varianz aus.

5.2 Ergebnisse der Bedienoberflächenstudie

Nachdem die Methode vorgestellt und ein Überblick über die relevanten Aspekte der Bedienoberflächenstudie gegeben wurde, erfolgt im folgenden Kapitel die Vorstellung der Ergebnisse des Experiments.

Zuerst werden die Ergebnisse bezogen auf den Einfluss der farblichen Gestaltung abgebildet. Danach erfolgt die Darstellung der Ergebnisse bezogen auf den Einfluss des Geschlechts der Versuchsteilnehmer und im Anschluss daran werden mögliche Abhängigkeiten zwischen den eben genannten Komplexen untersucht. Zum Abschluss dieses Kapitels werden die Ergebnisse zusammengefasst und kritisch gewürdigt.

5.2.1 Einfluss der farblichen Gestaltung

Um einen Überblick über die herangezogene Stichprobe zu geben, wird diese bei der Vorstellung der einzelnen abhängigen Variablen zu Beginn kurz dargestellt. Hierzu zählen neben der Stichprobengröße, auch die Ergebnisse der Varianzuntersuchung und des Tests auf Normalverteilung.

5.2.1.1 Ergebnisse bezogen auf die User Experience

Als Datengrundlage standen $N = 52$ Versuchsteilnehmer zur Verfügung. Die Varianzen waren homogen und die Daten der Dimensionen *Effizienz*, *Steuerbarkeit* und *Originalität* normalverteilt (Anhang G1).

Die Items des UEQ (Kapitel 5.1.3.2) wurden entweder von negativ zu positiv oder von positiv zu negativ bewertet. Da diese Ausrichtung zufällig gewählt war, mussten die Daten, bevor sie verwendet werden konnten, transformiert werden. Für die Umwandlung wurde ein, von den Entwicklern des UEQ erstelltes, Excel-Tool verwendet [NN15]. Nach der Umwandlung ergab sich ein Wertebereich der Items von -3 bis $+3$. Hierbei stand ein Wert von $+3$ für ein sehr positives Ergebnis, ein Wert von 0 für ein neutrales Ergebnis und ein Wert von -3 für ein sehr negatives Ergebnis.⁷⁵

5.2.1.1.1 User Experience der Bedienoberflächen (sechs Dimensionen des UEQ)

Zu Beginn wurde der Einfluss der farblichen Gestaltung auf die sechs Dimensionen des UEQ untersucht (Abbildung 5-11). Die Ergebnisse zeigen, dass in der Gruppe mit der dunklen Bedienoberfläche höhere Werte in vier der sechs Dimensionen gemessen wurden. Die helle

⁷⁵ Der äußere Wertebereich wird selten erreicht [RSP+13]. Erfahrungsgemäß bewegen sich die Mittelwerte durch Antworteffekte, wie z. B. die Vermeidung von Extremen, im Bereich von -2 bis $+2$. Bessere Einschätzungen werden nur selten erreicht, so dass ein Wert in der Nähe von $+/-2$ einem sehr positiven bzw. sehr negativen, nahezu optimalen Eindruck der Versuchsteilnehmer, entspricht.

Bedienoberfläche führte lediglich in den Dimensionen *Steuerbarkeit* und *Stimulation* zu einer besseren Einschätzung durch die Versuchsteilnehmer. Weiter auffällig war die Dimension *Originalität*. Im Vergleich zu den anderen Dimensionen wurden die Bedienoberflächen in dieser Dimension sehr schlecht bewertet. In der Gruppe mit der hellen Bedienoberfläche zeigte sich sogar eine negative Bewertung.

Um einen möglichen Einfluss von Ausreißern bzw. widersprüchlichen Antworten auf die Dimension *Originalität* ausschließen zu können, wurden die Antworten der Versuchsteilnehmer und die Skala auf ihre Konsistenz überprüft.⁷⁶ Hierbei wurden keine Ausreißer bzw. Unstimmigkeiten festgestellt. Die Bewertung der Originalität der Bedienoberflächen durch die Versuchsteilnehmer war homogen. Unabhängig davon, konnte ein Einfluss des Geschlechts der Versuchsteilnehmer zu diesem Zeitpunkt noch nicht ausgeschlossen werden.

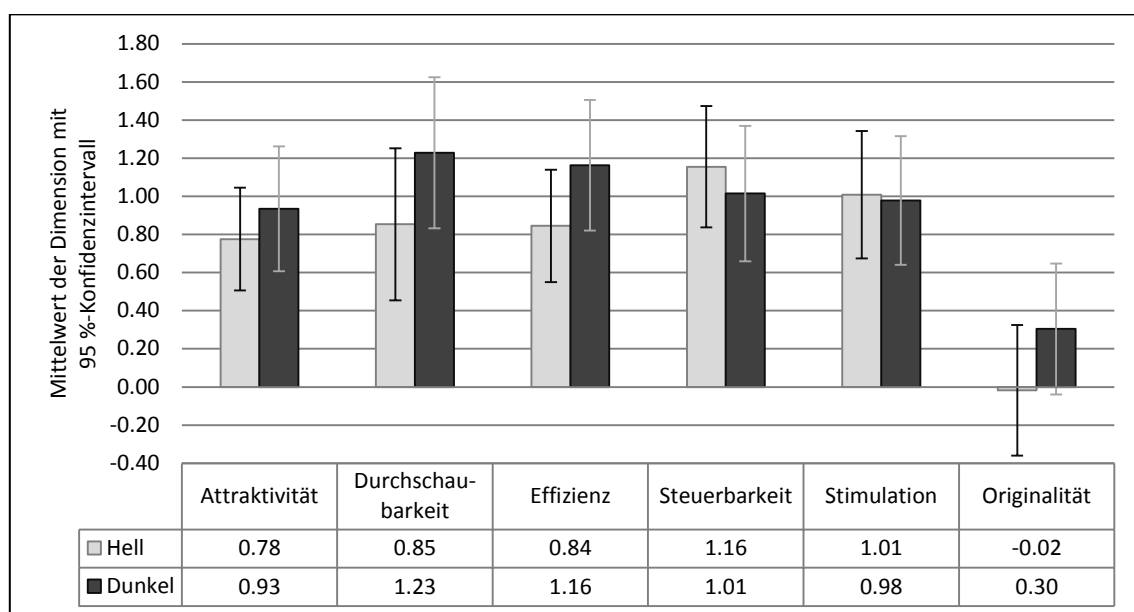


Abbildung 5-11: Mittelwerte und 95 %-Konfidenzintervalle der sechs Dimensionen bezogen auf die farbliche Gestaltung

Die beschriebenen Unterschiede konnten nicht als signifikant nachgewiesen werden (Anhang G1) und der Vergleich der Konfidenzintervalle⁷⁷ zeigte ebenfalls keine Auffälligkeiten.

5.2.1.1.2 User Experience der Bedienoberflächen (UX-Qualitäten nach HASSENZAHL)

Im nächsten Schritt wurden die Ergebnisse aus Abbildung 5-11 in die drei Qualitäten nach HASSENZAHL überführt (Kapitel 2.2.2.2.1). Hierfür wurde der Gesamtmittelwert über die Einzelmittelwerte der entsprechenden Dimensionen berechnet. Die pragmatische Qualität er-

⁷⁶ Für beide Analysen wurde das „mitgelieferte“ Excel-Tool der Entwickler verwendet [NN15].

⁷⁷ Das Konfidenzintervall ermöglicht eine erste grafische Beurteilung der Ergebnisse. Bei sich überlappenden Konfidenzintervallen, kann davon ausgegangen werden, dass die Unterschiede, bei $\alpha = .05$, nicht signifikant sind. Gleiches gilt, wenn der Wert 0 ein Teil des Konfidenzintervalls ist. Die Größe des Intervalls wird von dem Stichprobenumfang und der Streuung der Werte beeinflusst.

rechnete sich aus den Mittelwerten der Dimensionen *Durchschaubarkeit*, *Effizienz* und *Steuerbarkeit*. Die Dimensionen *Stimulation* und *Originalität* ergaben die hedonische Qualität. Die Attraktivität galt als globale Dimension und konnte deshalb direkt ausgewertet werden. Abbildung 5-12 zeigt die Ergebnisse der drei UX-Qualitäten bezogen auf den Einfluss der farblichen Gestaltung der Bedienoberfläche.

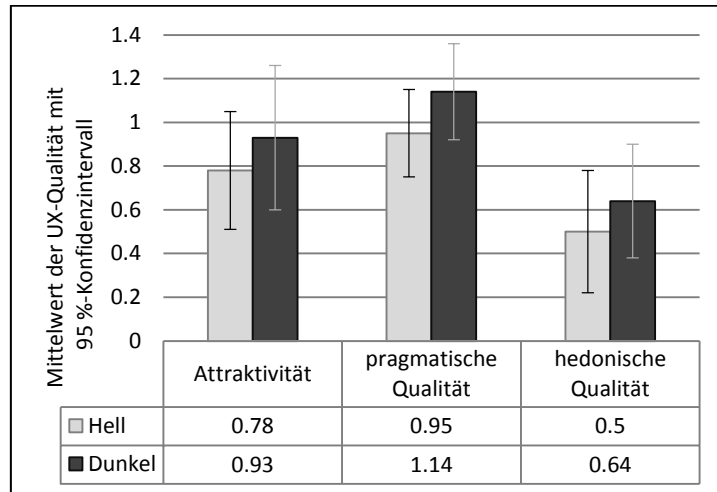


Abbildung 5-12: Mittelwerte der UX-Qualitäten Attraktivität, PQ und HQ bezogen auf die farbliche Gestaltung

Die dunkle Bedienoberfläche erzielte in jeder UX-Qualität eine bessere Bewertung als die helle Bedienoberfläche. Bedingt durch die schlechte Bewertung der Dimension *Originalität*, fiel die gemessene hedonische Qualität beider Bedienoberflächen entsprechend gering aus. Positiv an dieser Stelle war, dass der erfasste Unterschied die Tendenz aus der Voruntersuchung bestärkte. In Bezug auf die UX wurde die dunkle Bedienoberfläche besser bewertet als die helle Bedienoberfläche. Die Ergebnisse konnten jedoch nicht statistisch bestätigt werden, d. h. die Unterschiede zwischen den Gruppen waren nicht signifikant.

5.2.1.2 Ergebnisse bezogen auf die psychische Beanspruchung

Für die Auswertung der psychischen Beanspruchung standen $N = 50$ Versuchsteilnehmer zur Verfügung. Zwei Teilnehmer konnten nicht herangezogen werden, da sie den Fragebogen nicht korrekt ausgefüllt hatten. Die Daten des Items 1, des Items 2 und des Items 7 waren normalverteilt und die Varianzen waren homogen (Anhang G2).

Die Abbildung 5-13 zeigt die Mittelwerte und die 95 %-Konfidenzintervalle aller Items des Fragebogens zur Beurteilung des Trainings (Kapitel 5.1.3.3.2), die mit einer Ratingskala erfasst wurden. Die Ergebnisse der Freitextitems 8 und 9 werden gesondert in Tabelle 5-9 (Item 8) und in Tabelle 5-10 (Item 9) dargestellt.

Der Vergleich der Gruppen zeigte Unterschiede in den erfassten Zieldimensionen *Beanspruchung* und *Attraktivität*⁷⁸. Die Versuchsteilnehmer in der Gruppe mit der hellen Bedienoberfläche bewerteten die Beanspruchung während des Erlernens der Simulation (Item 1–3 und Item 7) größer als die Teilnehmer, welche die dunkle Bedienoberfläche verwendeten. Bei den Items, die eine Aussage über die Attraktivität der Bedienoberflächen zuließen (Item 4 und Item 5), verhielt es sich umgekehrt. Die farbliche Gestaltung und der Kontrast der dunklen Bedienoberfläche wurden von den Versuchsteilnehmern besser bewertet. Zudem erhielt die dunkle Bedienoberfläche bei der Fehlererkennung (Item 6) eine bessere Bewertung.

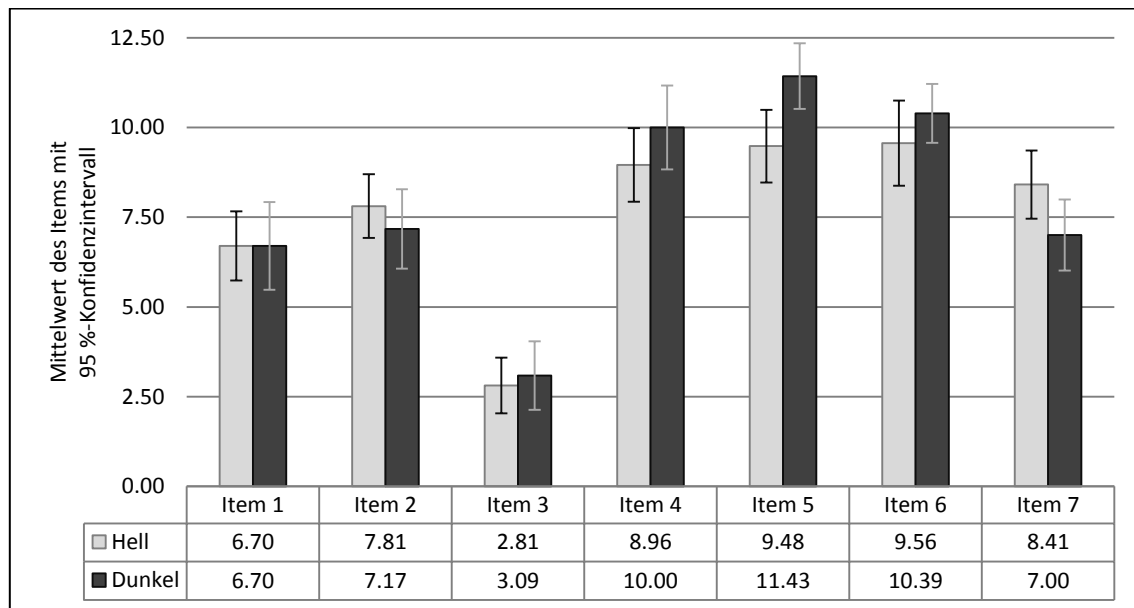


Abbildung 5-13: Mittelwerte und 95 %-Konfidenzintervalle der sieben Items bezogen auf die farbliche Gestaltung

Auffällig war auch, dass die Konfidenzintervalle des Items 5 nicht überlappten. Dies ließ auf einen signifikanten Unterschied schließen. Um dies zu überprüfen und um mögliche weitere signifikante Unterschiede feststellen zu können, wurden die Ergebnisse der Items im Einzelnen verglichen. Dabei werden die getrennt erfassten Dimensionen *Beanspruchung* und *Attraktivität* gesondert beschrieben.

5.2.1.2.1 Zieldimension *Beanspruchung*

Die Beanspruchung wurde mit den Items 1, 2, 3, 7 und dem Item 8 erfasst. Bei den ersten vier Items handelte es sich um Items, welche mit einer Ratingskala erfasst wurden. Item 8 war ein Freitextitem und wurde gesondert erfasst.

⁷⁸ Eine Übersicht der Items mit den erfassten Zielvariablen zeigt Tabelle 5-1 in Kapitel 5.1.3.3.2.

Item 1

Die Anstrengung, die für das Erlernen der Simulation benötigt wurde, wurde von den Versuchsteilnehmern für beide Bedienoberflächen annähernd gleich bewertet [Hell ($M = 6.70$, $SD = 2.84$); Dunkel ($M = 6.70$, $SD = 2.43$)]. Demnach konnte auch kein signifikanter Unterschied festgestellt werden (Anhang G2).

Item 2

Die geistigen Anforderungen während des Erlernens, wurden in der Gruppe mit der hellen Bedienoberfläche höher empfunden als in der Gruppe mit der dunklen Bedienoberfläche [Hell ($M = 7.81$, $SD = 2.24$); Dunkel ($M = 7.17$, $SD = 2.55$)]. Dieser Unterschied konnte nicht als signifikant nachgewiesen werden (Anhang G2).

Item 3

Die körperlichen Anforderungen während des Erlernens, wurden in der Gruppe mit der hellen Bedienoberfläche niedriger empfunden als in der Gruppe mit der dunklen Bedienoberfläche [Hell ($M = 2.81$, $SD = 1.96$); Dunkel ($M = 3.09$, $SD = 2.21$)]. Dieser Unterschied konnte nicht als signifikant nachgewiesen werden (Anhang G2).

Item 7

Dieses Item diente ebenfalls der Erfassung der Anstrengung. Es wurde verwendet, um eine mögliche Antworttendenz der Versuchsteilnehmer auf dem Item 1 zu kontrollieren (Kapitel 5.1.3.3.2). Im Gegensatz zu Item 1 wurde bei Item 7 ein Unterschied festgestellt. Die Anstrengung, die notwendig war um die Bedienung zu erlernen, wurde in der Gruppe mit der hellen Bedienoberfläche größer eingeschätzt als in der Gruppe mit der dunklen Bedienoberfläche [Hell ($M = 8.41$, $SD = 2.41$); Dunkel ($M = 7.00$, $SD = 2.30$)]. Dieser Unterschied stellte sich, sowohl mit dem t-Test ($t(48) = -2.105$, $p = .041$) als auch mit dem Mann-Whitney-U-Test ($U(27, 23) = 199$, $p = .028$), als signifikant heraus.

Item 8

Da nicht alle Versuchsteilnehmer die Möglichkeit in Anspruch nahmen das Freitextitem 8 für die Bewertung der Beanspruchung zu nutzen, stand nur eine begrenzte Anzahl an Antworten zur Verfügung. Aus diesem Grund fand keine statistische Auswertung statt und es wurde nur ein qualitativer Vergleich durchgeführt.

Für die Auswertung wurden die verschiedenen Antwortmöglichkeiten zuerst kategorisiert. Es ergaben sich fünf Antwortkategorien. Die Antwortkategorie und die Anzahl der gegebenen Antworten werden nachfolgend in Tabelle 5-9 dargestellt.

Tabelle 5-9: Übersicht über die Antwortkategorie des Freitextitems 8 und die Anzahl der Nennungen

Antwort	Anzahl der Nennungen	
	Hell (N = 23)	Dunkel (N = 19)
a) Erkennen von Unregelmäßigkeiten	14	6
b) Überblick über Betriebssituation	3	4
c) Arbeitsbelastung	8	4
d) Abkürzungen verstehen	4	6
e) Bedienung der Kommandofreigabe	1	2

Antwortkategorie a)

In diese Kategorie wurden die Antworten einsortiert, die das Erkennen und Identifizieren der Unregelmäßigkeiten thematisierten. Die Versuchsteilnehmer fanden die Diversität der Unregelmäßigkeiten sehr anstrengend. Gerade die Gleisfreimeldestörung wurde als sehr schwer zu erkennen eingestuft.

Antwortkategorie b)

Zu dieser Kategorie zählten die Antworten, die sich mit dem Überblick über die Betriebssituation befassten. Die Versuchsteilnehmer empfanden es als anstrengend den Gesamtüberblick zu behalten.

Antwortkategorie c)

In dieser Kategorie wurden die Antworten zusammengefasst, die sich mit der Arbeitsbelastung befassten. Die Versuchsteilnehmer bemängelten die vielen blinkenden Elemente und die Menge an parallel durchzuführenden Handlungen. Gerade in der Rückfallebene war die empfundene Beanspruchung hoch, da hier der Grad der technischen Unterstützung geringer war.

Antwortkategorie d)

Hier wurden alle Antworten aufsummiert, die sich mit dem Verständnis der angezeigten Informationen befassten. Die Versuchsteilnehmer fanden die vielen Abkürzungen, die für die Bedienung der Simulation notwendig waren, undurchsichtig und belastend.

Antwortkategorie e)

Zu dieser Kategorie zählten die Antworten, die sich mit dem Thema *Kommandofreigabe* befassten. Die Beachtung der Kommandofreigabe wurde als belastend empfunden. Vor allem die Lage auf dem Bildschirm und die umständliche Bedienung wurden hier benannt.

Die Lage am unteren Bildschirmrand wurde als störend empfunden, weil die Bedienung der Kommandofreigabe von der eigentlichen Betriebssituation ablenkte. Zudem wurde der Handlungsfluss durch die zusätzlichen Prüfaufgaben und die voreingestellte Ablaufzeit unterbrochen, was ebenfalls als belastend empfunden wurde.

An dieser Stelle werden die gewonnenen Erkenntnisse aus Tabelle 5-9 kurz zusammengefasst. In den Kategorien *Erkennen von Unregelmäßigkeiten* und *Arbeitsbelastung* war der Unterschied zwischen den Gruppen am größten. Insbesondere das Erkennen von Unregelmäßigkeiten wurde bei der hellen Bedienoberfläche als sehr belastend empfunden. Die Kategorie *Überblick über Betriebssituation* wurde von beiden Gruppen ähnlich bewertet. Wobei hier die Anzahl der Nennungen in der Gruppe mit der hellen Bedienoberfläche etwas größer war. Demnach konnte von einer leichten Bevorteilung der hellen Bedienoberfläche in dieser Kategorie ausgegangen werden. In den anderen Kategorien wurden keine aussagekräftigen Unterschiede festgestellt, zumal die weiteren Kategorien *Abkürzungen verstehen* und *Bedienung der Kommandofreigabe* vermutlich durch die fachfremde Stichprobe zu erklären sind. Die Verwendung von Abkürzungen im betrieblichen Umfeld und die Bedienung der Kommandofreigabe zählen zur täglichen Arbeit eines Fdl.

5.2.1.2.2 Zieldimension *Attraktivität*

Die Attraktivität wurde mit den Items 4, 5, 6 und dem Item 9 erfasst. Bei den ersten drei Items handelte es sich um Items, welche mit einer Ratingskala erfasst wurden. Item 9 war ein Freitextitem und wurde gesondert erfasst.

Item 4

Die farbliche Gestaltung der Bedienoberfläche wurde in der Gruppe mit der hellen Bedienoberfläche unangenehmer empfunden als in der Gruppe mit der dunklen Bedienoberfläche [Hell ($M = 8.96$, $SD = 2.59$); Dunkel ($M = 10.00$, $SD = 2.70$)]. Der Unterschied konnte mit dem t-Test als nicht signifikant ($t(48) = -1.383$, $p = .173$) und mit dem Mann-Whitney-U-Test als marginal signifikant ($U(27, 23) = 223$, $p = .085$) nachgewiesen werden. Da der Mann-Whitney-U-Test im vorliegenden Fall nur eine Tendenz anzeigte und alle Voraussetzungen für die Nutzung des t-Tests erfüllt waren, wurde das Ergebnis des t-Tests stärker gewichtet. Demnach konnte kein signifikanter Unterschied festgestellt werden.

Item 5

In der Gruppe mit der hellen Bedienoberfläche wurde der Kontrast unangenehmer empfunden als in der Gruppe mit der dunklen Bedienoberfläche [Hell ($M = 9.48$, $SD = 2.55$); Dunkel ($M = 11.43$, $SD = 2.11$)]. Dieser Unterschied stellte sich, sowohl mit dem t-Test ($t(48) = -2.922$, $p = .005$) als auch mit dem Mann-Whitney-U-Test ($U(27, 23) = 177.5$, $p = .009$), als signifikant heraus.

Item 6

Die Unterstützung durch die farbliche Gestaltung bei der Fehlererkennung wurde von den Versuchsteilnehmern bei der hellen Bedienoberfläche geringer eingeschätzt als bei der dunklen Bedienoberfläche [Hell ($M = 9.56$, $SD = 2.99$); Dunkel ($M = 10.39$, $SD = 1.90$)]. Dieser Unterschied konnte nicht als signifikant nachgewiesen werden (Anhang G2).

Item 9

Da nicht alle Versuchsteilnehmer die Möglichkeit in Anspruch nahmen das Freitextitem 9 für die Benennung von Darstellungsmängeln zu nutzen, stand nur eine begrenzte Anzahl an Antworten zur Verfügung. Analog zu Item 8 fand keine weitere statistische Auswertung statt.

Wie auch bei der Dimension *Beanspruchung* wurden die verschiedenen Antwortmöglichkeiten vorab kategorisiert. Im beschriebenen Fall ergaben sich fünf Antwortkategorien. Tabelle 5-10 zeigt nachfolgend die fünf Kategorien und die Anzahl der gegebenen Antworten.

Tabelle 5-10: Übersicht der Antwortkategorien des Freitextitems 9 und die Anzahl der Nennungen

Antwort	Anzahl der Nennungen	
	Hell ($N = 22$)	Dunkel ($N = 17$)
a) Informationskodierung	7	2
b) Ausleuchtung der Abschnittsbelegung	5	9
c) Lage der Kommandofreigabe	2	2
d) Kontrast der Elementbeschriftung	3	1
e) Hintergrundfarbe	4	1

Antwortkategorie a)

In dieser Kategorie wurden alle Antworten gesammelt, die sich mit der Informationskodierung befassten. Die Versuchsteilnehmer bemängelten die Anzahl und die Art der blinkenden Elemente. Zudem wurde die schlechte Unterscheidbarkeit von Regel- und Störzuständen als irritierend empfunden.

Antwortkategorie b)

Hierzu zählten die Antworten, die die schlechte Unterscheidbarkeit (bezogen auf die Farbgebung) zwischen Zugfahrt und Gleisfreimelde störung bemängelten. Die Versuchsteilnehmer fanden die rote Fahrwegausleuchtung störend und unübersichtlich. Der Hauptkritikpunkt der Teilnehmer war, dass die rote Ausleuchtung im Fahrweg ausschließlich für Störungen verwendet werden sollte und nicht für Zugfahrten.

Antwortkategorie c)

Wie auch bei dem Freitextitem 8, wurde die Darstellung der Kommandofreigabe als störend empfunden. In dieser Kategorie wurden die Antworten gesammelt, welche die Darstellung und die Lage der Kommandofreigabe auf dem Bildschirm bemängelten.

Antwortkategorie d)

Zu dieser Kategorie zählten alle Antworten, die sich mit dem Kontrast der Elementbeschriftung befassten. Die Versuchsteilnehmer fanden den Kontrast zwischen Schrift und Hintergrund zu gering und die Beschriftung teilweise zu dünn. Dadurch waren die Informationen zu unauffällig und konnten schlecht erkannt werden.

Antwortkategorie e)

In dieser Kategorie wurden die Antworten aufsummiert, die sich mit der Farbgebung des Hintergrunds beider Bedienoberflächen befassten. Die Versuchsteilnehmer fanden die Hintergrundfarbe zu hell bzw. zu dunkel.

An dieser Stelle werden die gewonnenen Erkenntnisse aus Tabelle 5-10 kurz zusammengefasst. In fast allen Kategorien wurde ein Unterschied zwischen den Gruppen festgestellt. Insbesondere die Kodierung der Informationen mittels Blinken, die Kontrastwirkung und die Hintergrundfarbe wurden bei der hellen Bedienoberfläche als irritierender und störender empfunden. In der Gruppe mit der dunklen Bedienoberfläche wurde die schlechte Unterscheidbarkeit zwischen Zugfahrt und Gleisfreimelde störung bemängelt. Die Versuchsteilnehmer konstatierten, dass die rote Ausleuchtung einer Zugfahrt zu einer schlechteren Erkennbarkeit der Gleisfreimelde störung führte. In der Kategorie *Lage der Kommandofreigabe* wurden keine aussagekräftigen Unterschiede festgestellt.

5.2.1.3 Ergebnisse bezogen auf die Befindlichkeit

Als Datengrundlage für die Auswertung der Befindlichkeit standen $N = 52$ Versuchsteilnehmer zur Verfügung. Die Daten waren nicht normalverteilt, aber die Varianzen waren homogen (Anhang G3).

Zu Beginn der Auswertung wurden die Differenzen der Vorher-Nachher-Messung der verschiedenen Dimensionen⁷⁹ gebildet und als vorläufiges Maß für den Einfluss der farblichen Gestaltung auf die Befindlichkeit genutzt (Abbildung 5-14).

⁷⁹ Für eine Übersicht und eine Erklärung der Dimensionen des MDBF siehe Tabelle 5-3 in Kapitel 5.1.3.4.2.

Der Vergleich der Differenzen zeigte deutlich, dass die helle Bedienoberfläche alle Dimensionen negativ beeinflusste. Die Befindlichkeit der Versuchsteilnehmer in der Gruppe mit der hellen Bedienoberfläche hatte sich nach der Eingewöhnungsphase verschlechtert. Im Gegensatz dazu, beeinflusste die dunkle Bedienoberfläche die Dimension *GS* kaum und die Dimension *WM* positiv. Die positive Veränderung in der Dimension *WM* zeigte eine Aktivierung der Versuchsteilnehmer. Ein weiterer Unterschied war, dass in der Gruppe mit der dunklen Bedienoberfläche eine stärker wachsende Unruhe festgestellt wurde. Dieser Effekt war zudem signifikant (Vorher-Nachher-Messung).

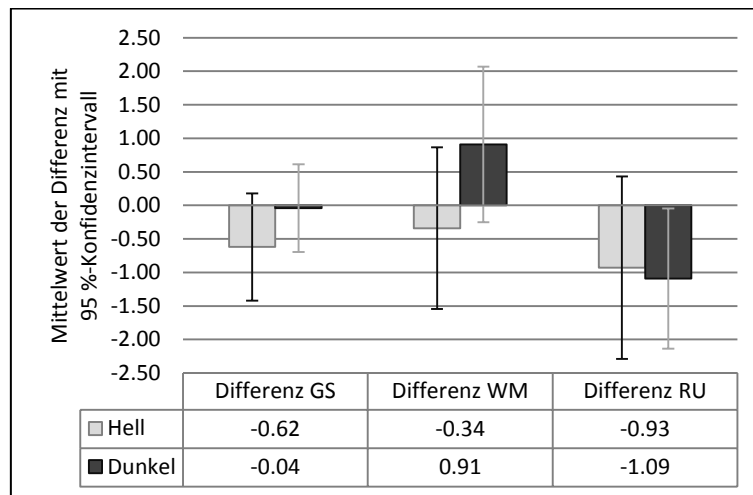


Abbildung 5-14: Mittelwerte und 95 %-Konfidenzintervalle der Differenzen der Dimensionen *Gute-Schlechte-Stimmung* (GS), *Wachheit-Müdigkeit* (WM) und *Ruhe-Unruhe* (RU) bezogen auf die farbliche Gestaltung

5.2.1.3.1 Zieldimension *Gute-Schlechte-Stimmung* (GS)

Um die dargestellten Ergebnisse besser nachvollziehen zu können, wird an dieser Stelle darauf hingewiesen, dass ein höherer Wert in der Nachher-Messung grundsätzlich für eine positive Entwicklung der betrachteten Dimension steht. Das heißt, dass sich die gemessene Dimension im Vergleich zur Eingangsmessung (Vorher-Messung) verbessert hat. Das gilt sowohl für den positiven als auch für den negativen Aspekt der Dimension.⁸⁰ Ein hoher Wert in der Nachher-Messung eines negativen Aspekts steht für das nicht Vorhandensein eben dieses.

Abbildung 5-15 verdeutlicht noch einmal die Gesamtveränderung der Dimension *GS*. In der Gruppe mit der hellen Bedienoberfläche war der negative Einfluss auf die Dimension *GS* stärker. Der Einfluss war jedoch ausschließlich qualitativer Natur, d. h. es konnte kein signifikanter Unterschied festgestellt werden (Anhang G3).

⁸⁰ Wie in 5.1.3.4.2 gezeigt wurde, setzen sich die beiden Kurzformen des MDBF aus zwölf Items zusammen. Jede Kurzform erfasst drei Dimensionen der Befindlichkeit. Dabei wird jede Dimension mit vier Items gemessen, wobei diese aus jeweils zwei positiv und zwei negativ orientierten Items bestehen.

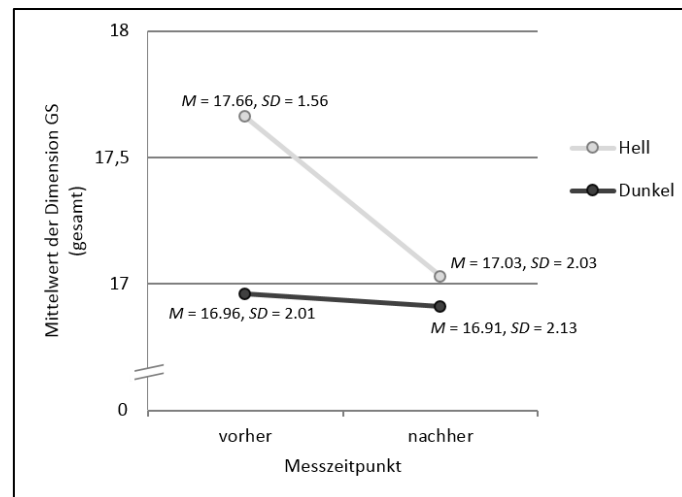


Abbildung 5-15: Mittelwerte und Standardabweichungen der Vorher-Nachher-Messung der Dimension *Gute-Schlechte-Stimmung* (GS) bezogen auf die farbliche Gestaltung

Auch wenn kein signifikanter Unterschied messbar war, wurde der Ursprung für die qualitative Abweichung untersucht. Dafür wurde die Gesamtveränderung der Dimension GS in ihren positiven und negativen Aspekt zerlegt, d. h. die Items wurden, wie in Kapitel 5.1.3.4.2 beschrieben, auf beide Aspekte aufgeteilt.

Der Vergleich des positiven Aspekts der Dimension GS ergab ein ähnliches Bild wie bei der Gesamtveränderung (Abbildung 5-16). Beide Bedienoberflächen beeinflussten den positiven Aspekt negativ. Dabei war die Veränderung bei beiden Bedienoberflächen annähernd gleich.

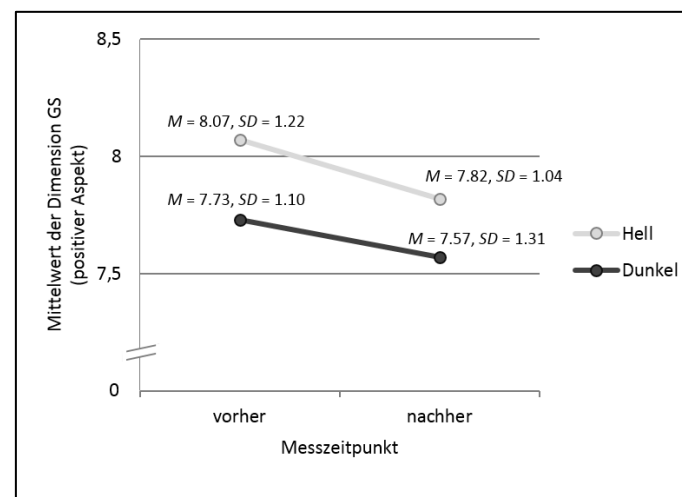


Abbildung 5-16: Mittelwerte und Standardabweichungen der Vorher-Nachher-Messung des positiven Aspekts der Dimension *Gute-Schlechte-Stimmung* (GS) bezogen auf die farbliche Gestaltung

Als nächstes wurde der negative Aspekt der Dimension GS untersucht (Abbildung 5-17). Hier offenbarte sich der Ursprung des Ergebnisses der Gesamtveränderung. Der negative Aspekt hatte sich bei der dunklen Bedienoberfläche verbessert. Das heißt, der negative Aspekt wurde durch die dunkle Bedienoberfläche abgeschwächt. Das war ein interessanter Unterschied zwischen den beiden Bedienoberflächen – wenn auch nur qualitativer Natur.

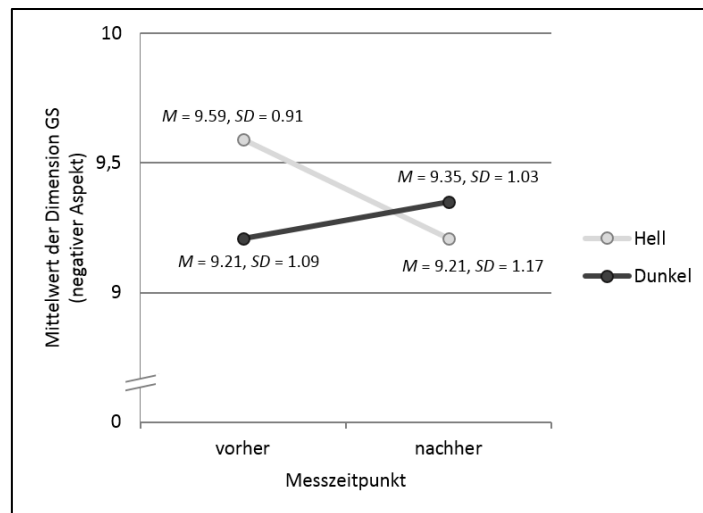


Abbildung 5-17: Mittelwerte und Standardabweichungen der Vorher-Nachher-Messung des negativen Aspekts der Dimension *Gute-Schlechte-Stimmung* (GS) bezogen auf die farbliche Gestaltung

Zusammenfassend beeinflussten beide Bedienoberflächen die Befindlichkeit in der Dimension GS negativ. Das zeigte sich auch bei dem positiven Aspekt. Im Gegensatz dazu, wurde der negative Aspekt der Dimension GS durch die dunkle Bedienoberfläche positiv beeinflusst. Das ist eine mögliche Erklärung dafür, dass die Gesamtveränderung etwas geringer ausfiel.

5.2.1.3.2 Zieldimension *Wachheit-Müdigkeit* (WM)

Abbildung 5-18 verdeutlicht die Gesamtveränderung. Der Vergleich der Bedienoberflächen zeigte in der Dimension WM einen weiteren Unterschied. Die farbliche Gestaltung der Bedienoberfläche beeinflusste die Wirkrichtung in dieser Dimension. Die dunkle Bedienoberfläche wirkte sich positiv auf die Versuchsteilnehmer aus. Sie waren nach der Eingewöhnungsphase wacher als die Teilnehmer in der Gruppe mit der hellen Bedienoberfläche. Wie auch in der Dimension GS war der Unterschied nicht signifikant (Anhang G3).

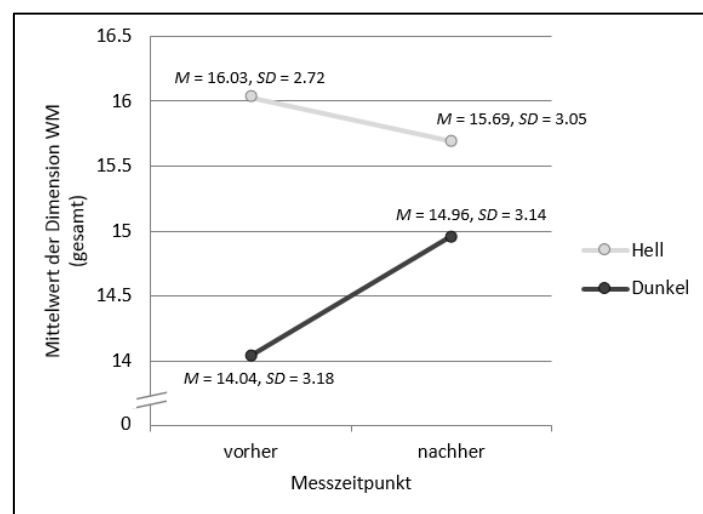


Abbildung 5-18: Mittelwerte und Standardabweichungen der Vorher-Nachher-Messung der Dimension *Wachheit-Müdigkeit* (WM) bezogen auf die farbliche Gestaltung

Der Vergleich des positiven Aspekts der Dimension *WM* ergab keinen Unterschied zur Gesamtveränderung. Abbildung 5-19 zeigt das Ergebnis. Die dunkle Bedienoberfläche führte zu einer Verbesserung des positiven Aspekts und die helle Bedienoberfläche zu einer Verschlechterung.

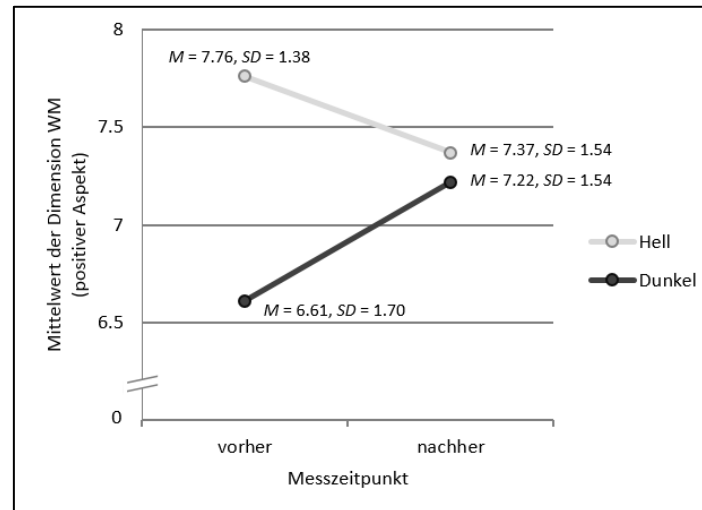


Abbildung 5-19: Mittelwerte und Standardabweichungen der Vorher-Nachher-Messung des positiven Aspekts der Dimension *Wachheit-Müdigkeit* (WM) bezogen auf die farbliche Gestaltung

Der Vergleich des negativen Aspekts der Dimension *WM* zeigte einen Einfluss beider Bedienoberflächen (Abbildung 5-20). Demnach wirkten sich beide Bedienoberflächen positiv auf den negativen Aspekt der Dimension *WM* aus.

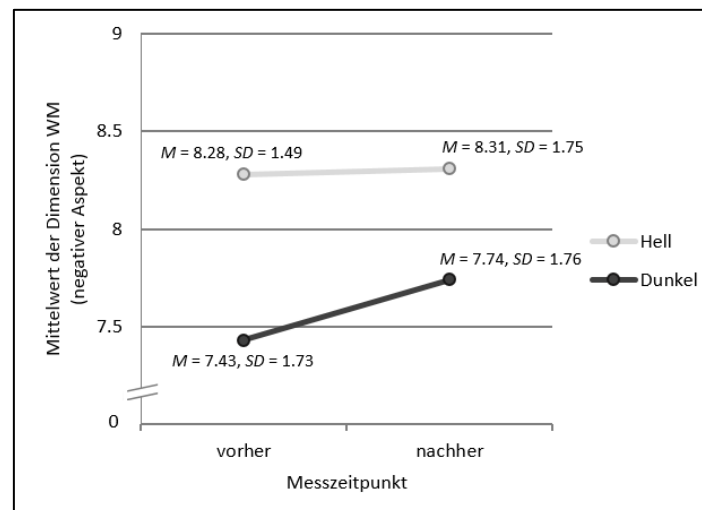


Abbildung 5-20: Mittelwerte und Standardabweichungen der Vorher-Nachher-Messung des negativen Aspekts der Dimension *Wachheit-Müdigkeit* (WM) bezogen auf die farbliche Gestaltung

Als Ergebnis für die Dimension *WM* kann festgehalten werden, dass beide Bedienoberflächen den negativen Aspekt der Dimension positiv beeinflussten. Hinzu kommt, dass die dunkle Bedienoberfläche, im Gegensatz zur hellen Bedienoberfläche, auch den positiven Aspekt der Dimension in positiver Richtung beeinflusste. Dadurch war der Einfluss der dunk-

len Bedienoberfläche insgesamt besser. Die Nutzung der dunklen Bedienoberfläche erhöhte die Wachheit und reduzierte gleichzeitig die Müdigkeit.

5.2.1.3.3 Zieldimension *Ruhe-Unruhe* (RU)

Auf den ersten Blick unterschieden sich die Bedienoberflächen in der Dimension *RU* nicht voneinander. Sowohl die helle als auch die dunkle Bedienoberfläche beeinflussten die Dimension negativ. Abbildung 5-21 verdeutlicht die Gesamtveränderung der Dimension *RU* für beide Bedienoberflächen. Im Gegensatz zu den Dimensionen *GS* und *WM* ergab sich in der Dimension *RU* ein signifikanter Unterschied. Dieser wurde bei der Vorher-Nachher-Messung in der Gruppe mit der dunklen Bedienoberfläche nachgewiesen. Er konnte sowohl mit dem t-Test für abhängige Stichproben ($t(22) = 2.163, p = .042$) als auch mit dem Wilcoxon-Test ($T(23) = 51, p = .024, z = -2.262$) nachgewiesen werden.⁸¹

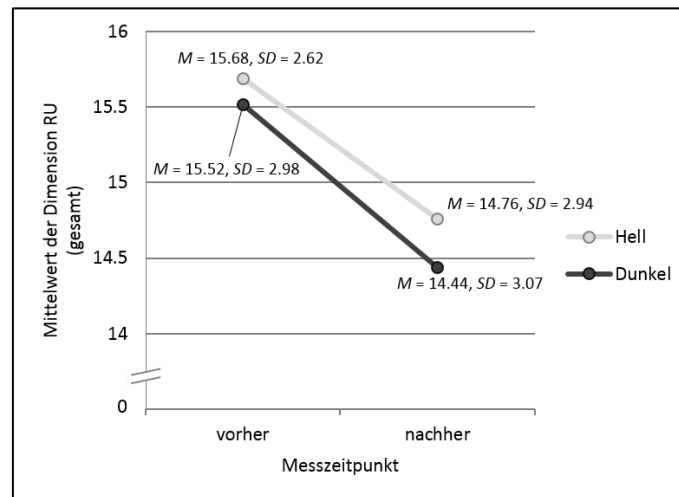


Abbildung 5-21: Mittelwerte und Standardabweichungen der Vorher-Nachher-Messung der Dimension *Ruhe-Unruhe* (RU) bezogen auf die farbliche Gestaltung

Um den Ursprung des signifikanten Unterschieds zu untersuchen, wurde die Gesamtveränderung erneut in den positiven und den negativen Aspekt zerlegt. Für den positiven Aspekt der Dimension *RU* ergab sich ein interessantes Bild (Abbildung 5-22). Beide Bedienoberflächen beeinflussten diesen Aspekt nur leicht bzw. gar nicht. Dies konnte demnach nicht der Ursprung des signifikanten Effekts sein.

⁸¹ Um einen möglichen Effekt zwischen den Ergebnissen der Vorher-Nachher-Messung zu untersuchen, wurde ein Test für Messwiederholung verwendet. Wie auch bei den anderen Untersuchungen kamen ein parametrisches und ein nicht parametrisches Verfahren zum Einsatz (Kapitel 5.1.6). Als parametrisches Verfahren wurde der t-Test für abhängige Stichproben verwendet und als nicht parametrisches Verfahren der Wilcoxon-Test.

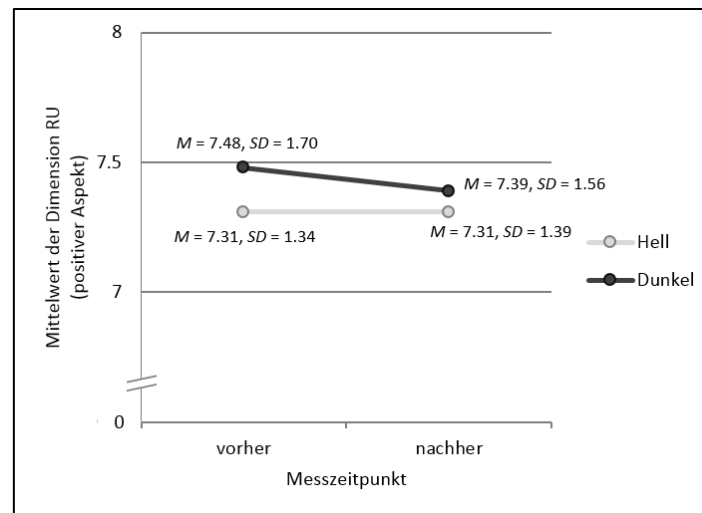


Abbildung 5-22: Mittelwerte und Standardabweichungen der Vorher-Nachher-Messung des positiven Aspekts der Dimension *Ruhe-Unruhe* (RU) bezogen auf die farbliche Gestaltung

Im nächsten Schritt wurde der negative Aspekt untersucht. Hier zeigte sich ebenfalls kein Unterschied (Abbildung 5-23). Beide Bedienoberflächen beeinflussten den negativen Aspekt der Dimension *RU* in negativer Wirkrichtung. Auch hier konnte der Ursprung des signifikanten Effekts nicht nachgewiesen werden.

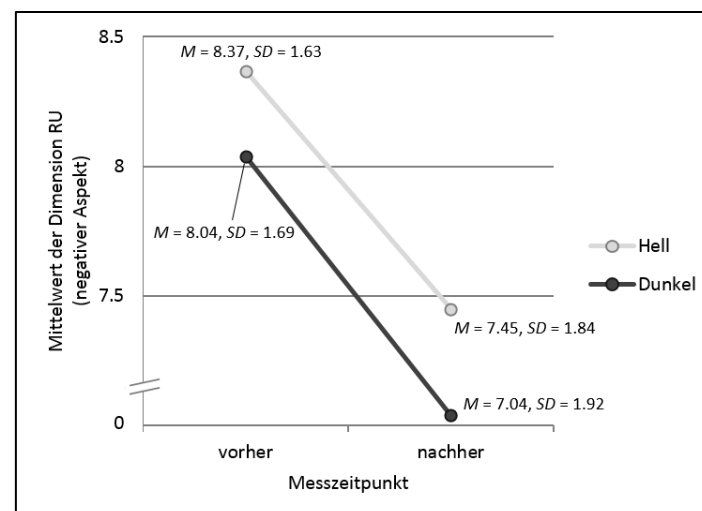


Abbildung 5-23: Mittelwerte und Standardabweichungen der Vorher-Nachher-Messung des negativen Aspekts der Dimension *Ruhe-Unruhe* (RU) bezogen auf die farbliche Gestaltung

Zusammenfassend beeinflussten beide Bedienoberflächen die Dimension *RU* negativ, wobei ein signifikanter Unterschied in der Gruppe mit der dunklen Bedienoberfläche nachgewiesen wurde. Dieser Effekt konnte, bezogen auf die farbliche Gestaltung, nicht nachvollzogen werden. Aus diesem Grund wurde ein geschlechterspezifischer Unterschied vermutet und im weiteren Verlauf der Arbeit untersucht (Kapitel 5.2.2).

5.2.1.4 Ergebnisse bezogen auf die Bedienungssicherheit

Als Datengrundlage wurden $N = 50$ Versuchsteilnehmer herangezogen. Zwei Teilnehmer konnten für die Auswertung nicht verwendet werden, da sie nicht für jeden Messpunkt eine Bewertung abgegeben hatten. Aus diesem Grund wurden sie bei der Betrachtung dieser abhängigen Variable komplett ausgeschlossen, um für alle Messpunkte die gleiche Anzahl an Antworten zu haben. Die Daten waren nicht normalverteilt, aber die Varianzen waren homogen (Anhang G4).

Wie in Kapitel 5.1.5.2 beschrieben, entwickelten sich der Automatisierungsgrad und die Schwierigkeit der zu erlernenden Unregelmäßigkeiten in der Eingewöhnungsphase entgegengesetzt. Anfangs war der Automatisierungsgrad niedrig und die Schwierigkeit hoch. Das Verhältnis änderte sich im Laufe der Eingewöhnungsphase. Durch dieses Vorgehen wurden die Versuchsteilnehmer zu Beginn der Eingewöhnungsphase stärker gefordert – was das Erlernen der manuellen Bedienungen unterstützen sollte – und später durch die technische Unterstützung entlastet – wodurch der Lernprozess gefördert werden sollte. Wie die Auswertung in Abbildung 5-24 zeigt, wurde dieses Ziel mit dem gewählten Verfahren erreicht. Die Versuchsteilnehmer fühlten sich mit steigendem Trainingsfortschritt sicherer. Statistisch konnte jedoch kein signifikanter Unterschied nachgewiesen werden (Anhang G4).

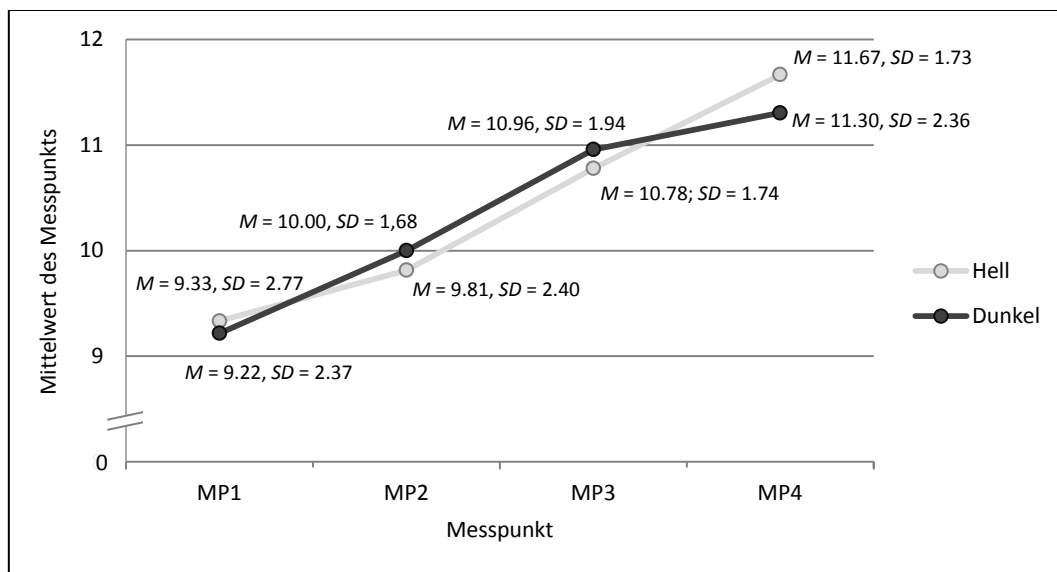


Abbildung 5-24: Mittelwerte und Standardabweichungen der vier Messpunkte (MP) für die Bedienungssicherheit bezogen auf die farbliche Gestaltung

Während der Eingewöhnungsphase wurde vom Versuchsleiter festgestellt, dass der größte Wissenszuwachs der Versuchsteilnehmer im dritten Trainingsblock (nach 40 min) zu erkennen war. Die Versuchsteilnehmer hatten zu diesem Zeitpunkt das Erlernte spürbar verinnerlicht und fühlten sich in der Bedienung sicher. Um diese subjektive Einschätzung des Versuchsleiters zu bestätigen, wurde die Differenz der Bewertungen zwischen den vier Messpunkten ermittelt und anschließend miteinander verglichen (Abbildung 5-25).

Anhand der Differenzen konnte die subjektive Einschätzung des Versuchsleiters bestätigt werden. Zwischen dem zweiten und dem dritten Messpunkt (MP3/MP2) war der Sicherheitszuwachs am größten.

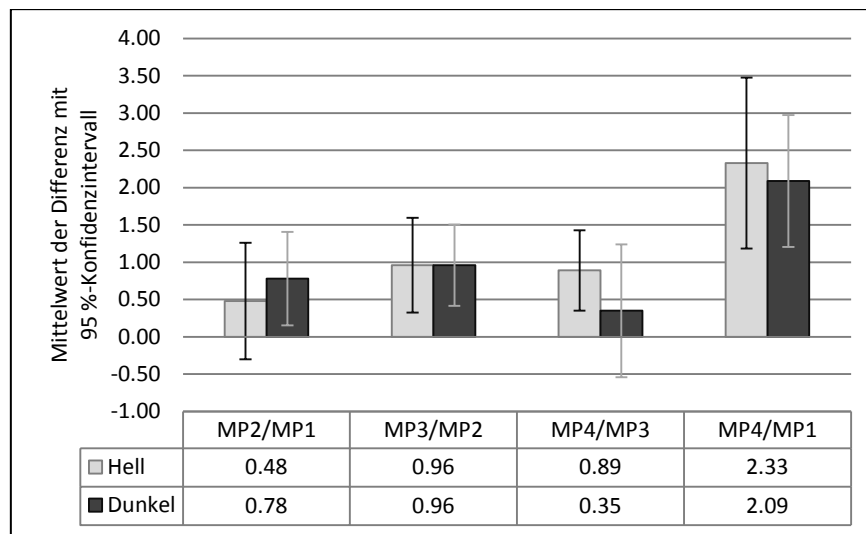


Abbildung 5-25: Mittelwerte und 95 %-Konfidenzintervalle der Messpunktdifferenzen bezogen auf die farbliche Gestaltung

Um zu überprüfen, ob die farbliche Gestaltung einen Einfluss auf den Sicherheitszuwachs über die gesamte Laufzeit hatte, wurde zusätzlich die Differenz zwischen dem ersten und dem letzten Messpunkt ermittelt (Abbildung 5-25). Der Gesamtzuwachs vom ersten bis zum letzten Messpunkt (MP4/MP1) war in der Gruppe mit der hellen Bedienoberfläche größer als in der Gruppe mit der dunklen Bedienoberfläche. [Hell ($M = 2.33$, $SD = 2.90$); Dunkel ($M = 2.09$, $SD = 2.04$)]. Dieser Unterschied war ausschließlich qualitativer Natur (Anhang G4).

5.2.1.5 Ergebnisse bezogen auf die vordefinierten Leistungskennzahlen

Für die Auswertung der Leistungskennzahlen standen 12 Versuchsteilnehmer nicht zur Verfügung. Die Ursachen waren zum einen technischer Natur und zum anderen ein unzureichender Trainingsstand der Teilnehmer⁸². Hiervon war hauptsächlich die Gruppe mit der hellen Bedienoberfläche betroffen.

Daraus ergab sich eine neue Stichprobenverteilung. Abbildung 5-26 zeigt die Zusammenstellung der beiden Versuchsgruppen nach der Reduzierung. Der Nachteil der neuen Verteilung war ein nicht ausgewogenes Verhältnis zwischen männlichen und weiblichen Versuchsteilnehmern in der Gruppe mit der hellen Bedienoberfläche. Um einen möglichen Einfluss dieses Umstands ausschließen zu können, wird in Kapitel 5.2.2 ebenfalls der Einfluss des Geschlechts der Versuchsteilnehmer auf die Leistungskennzahlen untersucht. Nachfolgend wird die neue Stichprobe im Teil vorgestellt.

⁸² Der Trainingsstand wurde mithilfe des standardisierten Szenarios (Testlauf) nach der Eingewöhnungsphase überprüft. Wenn die Versuchsteilnehmer ein Problem mit der manuellen Bedienung und/oder mit den Unregelmäßigkeiten hatten, wurden sie von der Auswertung ausgeschlossen.

Als Datengrundlage standen $N = 40$ Versuchsteilnehmer zur Verfügung. Diese setzten sich aus 38 Studierenden verschiedener Fachrichtungen, einem Angestellten der Universität und einer Privatperson zusammen. Die 17 weiblichen und 23 männlichen Versuchsteilnehmer waren zwischen 19 und 32 Jahre alt. Das Durchschnittsalter betrug 23.90 Jahre ($SD = 3.46$).

In der Gruppe mit der hellen Bedienoberfläche befanden sich 6 weibliche und 13 männliche Versuchsteilnehmer im Alter von 19 bis 31 Jahren ($M = 23.63$, $SD = 3.44$). Die Versuchsgruppe mit der dunklen Bedienoberfläche setzte sich aus 11 weiblichen und 10 männlichen Versuchsteilnehmern zusammen. Diese waren zwischen 19 und 32 Jahre alt. Das Durchschnittsalter betrug 24.14 Jahre ($SD = 3.54$). Im Vergleich zur Ausgangsstichprobe (Kapitel 5.1.1) ergab sich durch die Reduzierung eine Verringerung des Durchschnittsalters in der Versuchsgruppe mit der hellen Bedienoberfläche.

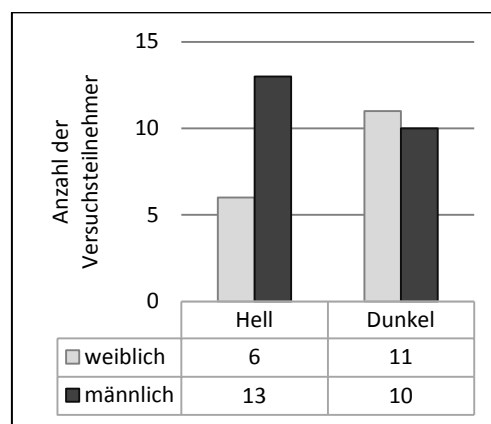


Abbildung 5-26: Aufteilung der Teilnehmer auf die Gruppen und nach Geschlecht (reduzierte Größe)

5.2.1.5.1 Anzahl der abgefertigten Zugfahrten

Die Anzahl der möglichen Zugfahrten während des 30-minütigen Hauptversuchs unterlag einer leichten Schwankung. Diese ergab sich aus nicht planbaren Störungen auf anderen Lupenbildern (Kapitel 4.3.3). Um am Ende der Studie die Anzahl der Zugfahrten vergleichen zu können und eine Aussage über die Leistungsfähigkeit der Versuchsteilnehmer treffen zu können, wurde vor der Auswertung die mittlere Anzahl der Zugfahrten bestimmt. Hierfür wurde der 30-minütige Hauptversuch⁸³ mit eingeschaltetem Zuglenkbetrieb und ohne Unregelmäßigkeiten einige Male ($N = 15$) aufgezeichnet. Anschließend wurde die Anzahl der Zugfahrten nach dem in Kapitel 5.1.4.2.1 beschriebenen Verfahren ermittelt. Die mittlere Anzahl der Zugfahrten betrug 52.33 Züge ($SD = 0.89$).

Die erfassten Daten für die Untersuchung der abgefertigten Zugfahrten waren nicht normalverteilt, aber die Varianzen waren homogen (Anhang G5).

⁸³ Da es eine gewisse Zeit in Anspruch nahm, bis sich ausreichend Züge im Netz der Stadtbahnsimulation befanden, wurde die Betriebssituation in der Simulation um 10 min „vorgespult“ (Kapitel 5.1.5.2).

In der Gruppe mit der hellen Bedienoberfläche wurden weniger Zugfahrten abgefertigt als in der Gruppe mit der dunklen Bedienoberfläche [Hell ($M = 45.05$, $SD = 6.69$); Dunkel ($M = 48.05$, $SD = 4.39$)]. Der Unterschied konnte, sowohl mit dem t-Test ($t(38) = -1.690$, $p = .099$) als auch mit dem Mann-Whitney-U-Test ($U(21, 19) = 133$, $p = .074$), als marginal signifikant nachgewiesen werden.

5.2.1.5.2 Fehlerentdeckungszeit

Die Zeiten, die für das Erkennen einer Unregelmäßigkeit benötigt wurden, schwankten sehr stark. Sie reichten von wenigen Sekunden, über mehrere Minuten bis hin zum Nichterkennen der Unregelmäßigkeit. Aus diesem Grund wurden die Daten vor der Auswertung bereinigt und aufbereitet. Dies geschah in drei Stufen.

In der ersten Stufe wurden die Ausreißer in den Einzelereignissen geglättet. Es gab Versuchsteilnehmer, welche für das Erkennen eines Einzelereignisses unverhältnismäßig lang brauchten. In diesem Fall wurde der Extremwert auf das Niveau des nächsthöheren Wertes in der Gruppe abgesenkt (konservative Betrachtung). Als Beispiel wurde der Ausfall des ZLB am S22 von einem Versuchsteilnehmer erst nach 415 s erkannt. Dieser Teilnehmer befand sich in der Gruppe mit der hellen Bedienoberfläche. Die nächsthöhere Entdeckungszeit in „seiner“ Gruppe betrug 188 s. Um nun den Ausreißer (415 s) nicht so stark in die Auswertung einzu beziehen, wurde die Entdeckungszeit dieses Versuchsteilnehmers auf den niedrigeren Wert (188 s) abgesenkt. Durch das Vorgehen wurden die Messwerte zusammengeführt und ein möglicher Trend blieb trotz der konservativen Betrachtung unberührt.

Als nächstes wurden die Unregelmäßigkeiten angepasst, die von den Versuchsteilnehmern gar nicht erkannt wurden (Stufe 2). Um diese Werte in die Auswertung einbeziehen zu können, wurde hierfür ein Maximalwert festgelegt. Nach der ersten Glättung (Stufe 1) näherte sich der Höchstwert der einzelnen Fehlerentdeckungszeiten einem Wert von 200 s an. Aus diesem Grund wurde für ein nicht erkanntes Ereignis die Entdeckungszeit auf 200 s gesetzt.

Im letzten Schritt wurden die Messwerte, die kleiner als eine Sekunde waren, bereinigt (Stufe 3). Wie in Kapitel 5.1.3.7 beschrieben, konnten die Fehlerentdeckungszeiten mit dem verwendeten Eyetracker nicht automatisiert ausgewertet werden. Um die Fehlerentdeckungszeiten dennoch erfassen zu können, wurden die Augenbewegungen beobachtet und die Zeiten manuell gestoppt. Das Problem an diesem Vorgehen war, dass die Fehlerentdeckungszeiten, die kleiner als eine Sekunde waren, nur geschätzt werden konnten. Um diese Messwerte für die Auswertung trotzdem verwenden zu können, wurde für sie ein fiktiver Wert von 500 ms definiert.⁸⁴

Nach der Aufbereitung der Daten wurden zuerst die Gesamtfehlerentdeckungszeiten miteinander verglichen. Hierfür wurden die Einzelfehlerentdeckungszeiten aufsummiert. Die Daten waren nicht normalverteilt, aber die Varianzen waren homogen (Anhang G5).

⁸⁴ Dabei handelt es sich um eine Festlegung des Versuchsleiters.

In der Gruppe mit der hellen Bedienoberfläche war die Gesamtfehlerentdeckungszeit größer als in der Gruppe mit der dunklen Bedienoberfläche. Das heißt, dass die Versuchsteilnehmer in der Gruppe mit der hellen Bedienoberfläche durchschnittlich länger brauchten, um eine Unregelmäßigkeit zu erkennen. [Hell ($M = 652.21$, $SD = 263.44$); Dunkel ($M = 628.45$, $SD = 270.02$)]. Dieser Unterschied konnte nicht als signifikant nachgewiesen werden (Anhang G5).

Die Untersuchung der Einzelereignisse ergab, bis auf ein paar Auffälligkeiten, keine aussagekräftigen Unterschiede. Abbildung 5-27 zeigt die Mittelwerte der Fehlerentdeckungszeiten der einzelnen Unregelmäßigkeiten grafisch und Tabelle 5-11 gibt einen Überblick über die Zahlenwerte. Die Bezeichnung der Unregelmäßigkeit in Abbildung 5-27 setzt sich aus der Ordnungsnummer, der Art der Unregelmäßigkeit und dem dazugehörigen Element zusammen. Dabei stehen die Akronyme *ZLB* für den Ausfall des Zuglenkbetriebs (Kapitel 4.3.2.1), *SS* für die Signalstörung (Kapitel 4.3.2.2) und *GFM* für die Gleisfreimeldestörung an einer Weiche (Kapitel 4.3.2.3). Im Folgenden werden nur die Unregelmäßigkeiten analysiert, bei denen der Unterschied zwischen den beiden Gruppen am aussagekräftigsten war.

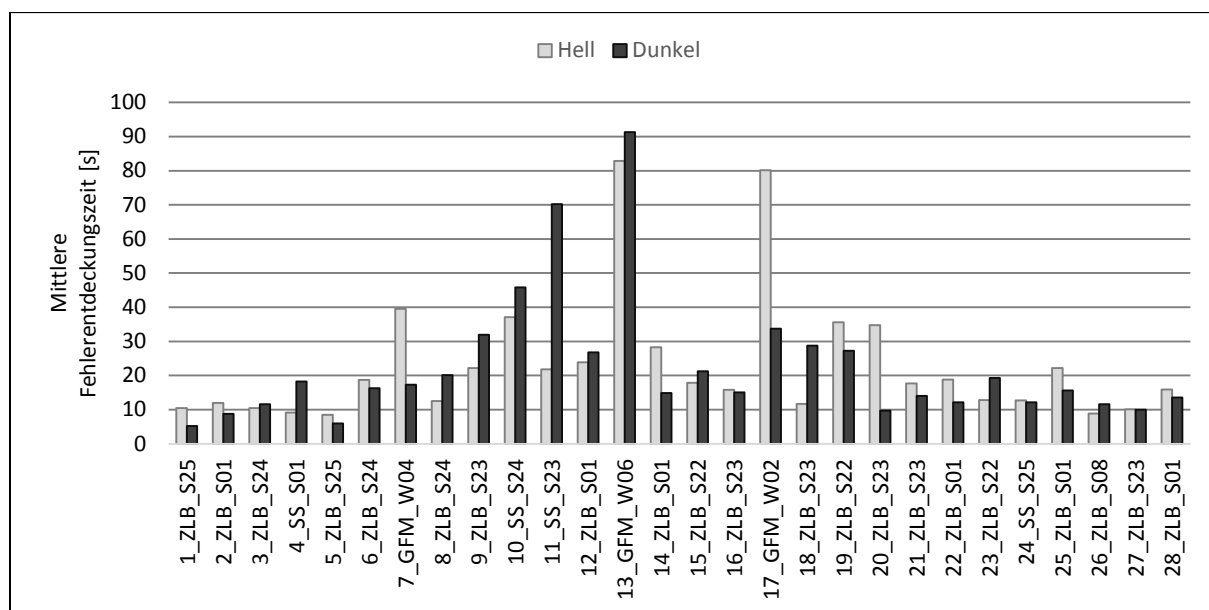


Abbildung 5-27: Mittelwerte der Fehlerentdeckungszeiten für die einzelnen Unregelmäßigkeiten bezogen auf die farbliche Gestaltung

Der Ursprung des Unterschieds bei der mittleren Fehlerentdeckungszeit der Unregelmäßigkeit *7_GFM_W04* konnte, bezogen auf die farbliche Gestaltung der Bedienoberfläche, nicht nachvollzogen werden. Beide Gruppen waren in gleichem Maße abgelenkt und die Unregelmäßigkeit wurde von allen Versuchsteilnehmern erkannt. An dieser Stelle wurde ein Einfluss des Geschlechts der Versuchsteilnehmer vermutet und deshalb in Kapitel 5.2.2.5.2 untersucht.

Der Unterschied zwischen den Gruppen bei der Fehlerentdeckungszeit der Unregelmäßigkeit *11_SS_S23* konnte nicht nachvollzogen werden. Auffällig war hier, dass das Erkennen der Signalstörung bei der Gruppe mit der dunklen Bedienoberfläche in drei von vier Fällen⁸⁵ länger gedauert hat. Zudem wurden die Unregelmäßigkeiten *10_SS_S24* und *11_SS_S23* in der Gruppe mit der dunklen Bedienoberfläche zweimal überhaupt nicht erkannt. Dieser Umstand und die zeitliche Einordnung – gerade am Anfang des Hauptversuchs – ließ die Vermutung zu, dass es sich hierbei um eine systematische Abweichung handelte.

Die hohen Fehlerentdeckungszeiten der Unregelmäßigkeit *13_GFM_W06* ließen sich durch die zeitliche Abfolge der Unregelmäßigkeiten erklären (Anhang D). Die Gleisfreimeldestörung der Weiche 06 trat direkt nach der Gleisfreimeldestörung der Weiche 04 auf. Durch die Störung der Weiche 04 kam es zu einem Rückstau von Zugfahrten. Da die Fahrstraße *S11–S20* betrieblich bedingt eine höhere Priorität in der einprogrammierten Zuglenkung hatte, wurden diese Fahrten von der Simulation zuerst abgefertigt. Dadurch verzögerte sich die Abfertigung der Zugfahrten über die Fahrstraße *S22–S20*. Das Problem daran war, dass zwischen dem Signal *S24* und dem Signal *S22* nur der Gleisfreimeldeabschnitt der Weiche 06 und der des Achszählers 242 lag. Stand ein Zug vor dem Signal *S22*, war der Gleisfreimeldeabschnitt des Achszählers 242 belegt. In diesem Fall mussten die Versuchsteilnehmer erkennen, dass sich der Zug nicht mehr bewegte und der Gleisfreimeldeabschnitt der Weiche 06 eigentlich frei sein musste. Für die richtige Erfassung der Betriebssituation in diesem Bereich, benötigten die ungeübteren Versuchsteilnehmer eine gewisse Zeit.

Die Abweichung bei der Fehlerentdeckungszeit der Unregelmäßigkeit *17_GFM_W02* rührte vermutlich daher, dass sie in der Gruppe mit der hellen Bedienoberfläche fünfmal nicht erkannt wurde. Da die männlichen Versuchsteilnehmer hierbei zu einem großen Anteil (viermal) beteiligt waren, wird in Kapitel 5.2.2.5.2 zusätzlich der Einfluss des Geschlechts untersucht. Eine weitere Erklärung für diesen Umstand war zudem, dass die Unregelmäßigkeit *17_GFM_W02* sehr schwer zu erkennen war – das galt unabhängig von der Gruppe. Der Grund hierfür war, dass die Gleisfreimeldestörung der Weiche 02 parallel zur Gleisfreimeldestörung der Weiche 06 eintrat. Die Versuchsteilnehmer wurden während dieses Zeitpunkts sehr stark belastet.

Der Ursprung des Unterschieds bei der mittleren Fehlerentdeckungszeit der Unregelmäßigkeit *18_ZLB_S23* konnte, bezogen auf die farbliche Gestaltung der Bedienoberfläche, nicht nachvollzogen werden. Beide Gruppen waren in gleichem Maße abgelenkt und die Unregelmäßigkeit wurde von allen Versuchsteilnehmern erkannt. An dieser Stelle wurde ein Einfluss des Geschlechts der Versuchsteilnehmer vermutet und deshalb in Kapitel 5.2.2.5.2 untersucht.

⁸⁵ Das Erkennen der Unregelmäßigkeiten *4_SS_S01*, *10_SS_S24* und *11_SS_S23* dauerte in der Gruppe mit der dunklen Bedienoberfläche länger als in der Gruppe mit der hellen Bedienoberfläche. Ausschließlich die Signalstörung *24_SS_S25* am Ende des Hauptversuchs wurde schneller erkannt.

Analog zu der Unregelmäßigkeit 18_ZLB_S23 konnte der Ursprung des Unterschieds bei der mittleren Fehlerentdeckungszeit der Unregelmäßigkeit 20_ZLB_S23 nicht nachvollzogen werden. Beide Gruppen waren in gleichem Maße abgelenkt und die Unregelmäßigkeit wurde von allen Versuchsteilnehmern erkannt. An dieser Stelle wurde ein Einfluss des Geschlechts der Versuchsteilnehmer vermutet und deshalb in Kapitel 5.2.2.5.2 untersucht.

Wie in Kapitel 5.1.5.3 beschrieben, wurde die Unregelmäßigkeit 26_ZLB_S08 eingeführt, um Defizite im Überwachungsverhalten der Versuchsteilnehmer festzustellen. Der Vergleich der Mittelwerte in den Gruppen zeigte keine Auffälligkeit. Beide Gruppen erkannten die Unregelmäßigkeit sehr schnell. Demnach konnte kein Effekt zwischen der farblichen Gestaltung und der dauernden Aufmerksamkeit festgestellt werden. Beide Gruppen waren am Ende des Hauptversuchs fähig ein seltenes Ereignis zu erkennen.

Tabelle 5-11: Mittelwerte und Standardabweichungen der Fehlerentdeckungszeiten für die einzelnen Unregelmäßigkeiten bezogen auf die farbliche Gestaltung

Nr.	Hell (N = 19)		Dunkel (N = 21)		Nr.	Hell (N = 19)		Dunkel (N = 21)	
	M	SD	M	SD		M	SD	M	SD
1_ZLB_S25	10.47	14.52	5.21	5.29	15_ZLB_S22	17.84	18.98	21.21	26.74
2_ZLB_S01	11.94	18.14	8.81	8.48	16_ZLB_S23	15.79	14.89	15.07	21.87
3_ZLB_S24	10.47	7.84	11.57	8.67	17_GFM_W02	80.13	79.52	33.69	37.10
4_SS_S01	9.11	8.17	18.24	22.25	18_ZLB_S23	11.66	12.12	28.74	31.70
5_ZLB_S25	8.52	8.82	5.98	4.60	19_ZLB_S22	35.55	56.43	27.26	30.54
6_ZLB_S24	18.71	25.66	16.29	20.85	20_ZLB_S23	34.71	47.85	9.71	10.78
7_GFM_W04	39.58	44.22	17.26	18.89	21_ZLB_S23	17.66	23.04	13.98	14.87
8_ZLB_S24	12.53	13.35	20.12	21.31	22_ZLB_S01	18.84	22.44	12.17	11.47
9_ZLB_S23	22.16	20.59	31.91	35.66	23_ZLB_S22	12.79	15.59	19.31	21.05
10_SS_S24	37.05	26.81	45.86	51.22	24_SS_S25	12.74	17.19	12.19	15.95
11_SS_S23	21.84	17.47	70.19	62.34	25_ZLB_S01	22.18	24.66	15.60	17.22
12_ZLB_S01	23.90	22.79	26.81	27.77	26_ZLB_S08	8.84	10.11	11.60	10.48
13_GFM_W06	82.90	65.22	91.33	67.20	27_ZLB_S23	10.13	12.83	9.95	14.11
14_ZLB_S01	28.26	23.17	14.88	14.48	28_ZLB_S01	15.90	18.58	13.52	9.42

5.2.1.5.3 Verhältnis zwischen Klickzahl und gemachten Fehlern in der Nebenaufgabe

Durch technische Probleme konnte ein Versuchsteilnehmer für die Auswertung nicht herangezogen werden. Die Daten waren nicht normalverteilt, aber die Varianzen waren homogen (Anhang G5).

In der Gruppe mit der hellen Bedienoberfläche wurde ein schlechteres Verhältnis zwischen der Klickzahl und den gemachten Fehlern beobachtet als in der Gruppe mit der dunklen Bedienoberfläche [Hell ($M = 107.16$, $SD = 103.39$); Dunkel ($M = 185.80$, $SD = 117.62$)]. Der Unterschied konnte, sowohl mit dem t-Test ($t(37) = -2.213$, $p = .033$) als auch mit dem Mann-Whitney-U-Test ($U(20, 19) = 94$, $p = .006$), als signifikant nachgewiesen werden.

5.2.1.5.4 Anzahl der Hilfshandlungen

Für die Auswertung wurde die Gesamtanzahl der Hilfshandlungen verwendet. Die erfassten Daten waren nicht normalverteilt, aber die Varianzen waren homogen (Anhang G5).

In der Gruppe mit der hellen Bedienoberfläche wurden weniger Hilfshandlungen durchgeführt als in der Gruppe mit der dunklen Bedienoberfläche [Hell ($M = 9.74$, $SD = 3.11$); Dunkel ($M = 11.05$, $SD = 2.38$)]. Dieser Unterschied konnte nicht als signifikant nachgewiesen werden (Anhang G5). Deshalb erfolgte keine weitere Untersuchung der Einzelbestandteile *VS-Fahrt Weiche* und *VS-Fahrt Signal*.

5.2.1.5.5 Anzahl der Bedienfehler

Für die Auswertung wurde die Gesamtanzahl der Bedienfehler verwendet. Die erfassten Daten waren nicht normalverteilt, aber die Varianzen waren homogen (Anhang G5).

Während des Hauptversuchs wurden in der Gruppe mit der hellen Bedienoberfläche mehr Bedienfehler gezählt als in der Gruppe mit der dunklen Bedienoberfläche [Hell ($M = 6.47$, $SD = 5.07$); Dunkel ($M = 4.38$, $SD = 3.76$)]. Dieser Unterschied war nicht signifikant (Anhang G5). Aus diesem Grund erfolgte keine weitere Untersuchung bezogen auf die Einzelklassifizierung der Bedienfehler.

5.2.1.5.6 Gesamtüberwachungszeit

Die erfassten Daten waren nicht normalverteilt und die Varianzen waren nicht homogen (Anhang G5).

In der Gruppe mit der hellen Bedienoberfläche wurde mehr Zeit mit der Überwachung des Bedienbereichs zugebracht als in der Gruppe mit der dunklen Bedienoberfläche [Hell ($M = 16.61$, $SD = 2.12$); Dunkel ($M = 15.51$, $SD = 3.52$)]. Dieser Unterschied konnte nicht als signifikant nachgewiesen werden (Anhang G5).

5.2.2 Einfluss des Geschlechts der Versuchsteilnehmer

Nachdem der Einfluss der farblichen Gestaltung untersucht und die Ergebnisse dargestellt wurden, wird im folgenden Kapitel der Einfluss des Geschlechts der Versuchsteilnehmer in der Bedienoberflächenstudie analysiert. Die Ergebnisse werden in gewohnter Weise vorgestellt.

Zu Beginn wird die Aufteilung der Stichprobe kurz dargestellt. Die weibliche Gruppe bestand aus 25 und die männliche Gruppe aus 27 Versuchsteilnehmern. Die weiblichen Versuchsteilnehmer waren zwischen 19 und 54 Jahre alt. Das Durchschnittsalter betrug 25.00 Jahre ($SD = 6.99$). Die männlichen Versuchsteilnehmer waren im Alter von 19 bis 52 Jahren ($M = 25.44$, $SD = 6.71$).

5.2.2.1 Ergebnisse bezogen auf die User Experience

Als Datengrundlage standen $N = 52$ Versuchsteilnehmer zur Verfügung. Die Varianzen waren homogen und die Daten der Dimensionen *Effizienz*, *Steuerbarkeit* und *Originalität* normalverteilt (Anhang G1).

5.2.2.1.1 User Experience der Bedienoberflächen (sechs Dimensionen des UEQ)

Analog zu Kapitel 5.2.1.1 wurde zu Beginn der Einfluss des Geschlechts auf die sechs Dimensionen des UEQ untersucht. Abbildung 5-28 zeigt die Mittelwerte und die 95 %-Konfidenzintervalle der transformierten Merkmalsausprägungen.

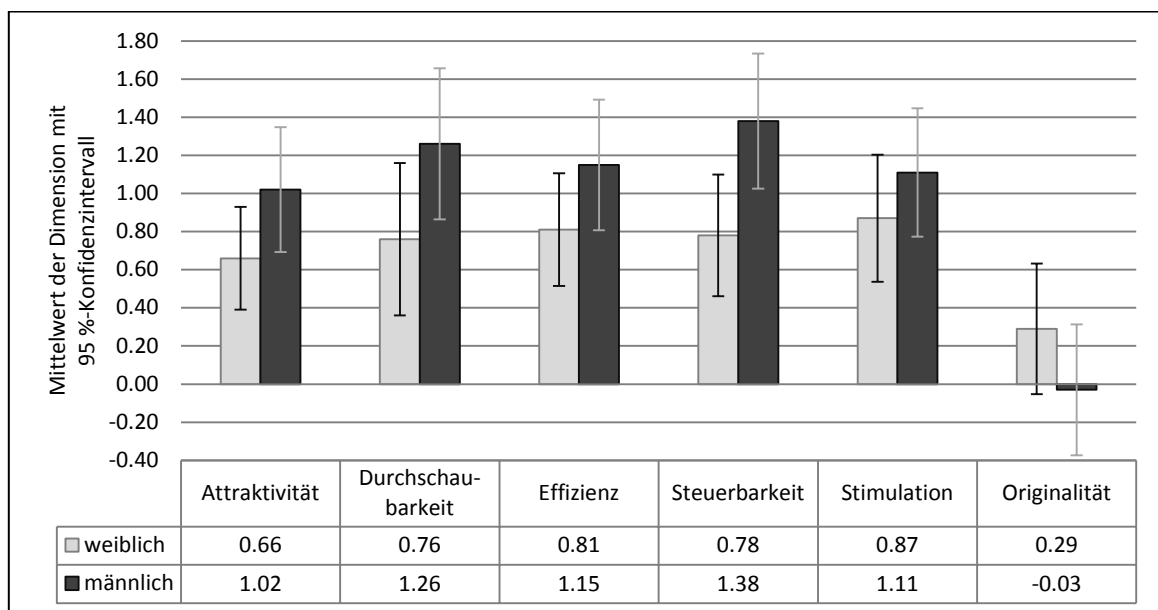


Abbildung 5-28: Mittelwerte und 95 %-Konfidenzintervalle der sechs Dimensionen bezogen auf das Geschlecht der Versuchsteilnehmer

Die männlichen Versuchsteilnehmer bewerteten die Dimensionen *Attraktivität*, *Durchschaubarkeit*, *Effizienz*, *Steuerbarkeit* und *Stimulation* der Bedienoberflächen besser als die weiblichen Versuchsteilnehmer. Demnach war ein Effekt, verursacht durch das Geschlecht der Versuchsteilnehmer, nachweisbar. Der gefundene Unterschied in der Dimension *Steuerbarkeit* war zudem signifikant. Er konnte sowohl mit dem t-Test ($t(50) = -2.620$, $p = .012$) als auch mit dem Mann-Whitney-U-Test ($U(29, 23) = 210.5$, $p = .019$) nachgewiesen werden.

Wie auch bei den Ergebnissen bezogen auf den Einfluss der farblichen Gestaltung, wurde die Dimension *Originalität* der Bedienoberflächen von beiden Gruppen sehr schlecht bewertet. Die männlichen Versuchsteilnehmer bewerteten die Dimension sogar negativ. Um zu überprüfen, ob die Ergebnisse der Dimension *Originalität* aus einer Wechselbeziehung zwischen farblicher Gestaltung und Geschlecht herrührten, wurde dies zusätzlich inferenzstatistisch untersucht (Kapitel 5.2.3.1).

5.2.2.1.2 User Experience der Bedienoberflächen (UX-Qualitäten nach HASSENZAHL)

Im Anschluss wurden die Ergebnisse in die drei Qualitäten nach *HASSENZAHL* überführt. Abbildung 5-29 zeigt die Ergebnisse bezogen auf das Geschlecht der Versuchsteilnehmer.

Die männlichen Versuchsteilnehmer bewerteten die Attraktivität und die pragmatische Qualität der Bedienoberflächen besser als die weiblichen Versuchsteilnehmer. Ausschließlich die hedonische Qualität der Bedienoberflächen wurde von den weiblichen Versuchsteilnehmern höher bewertet. Wie auch bei der farblichen Gestaltung, handelte es sich hierbei nur um qualitative Unterschiede, welche sich nicht als signifikant herausstellten.

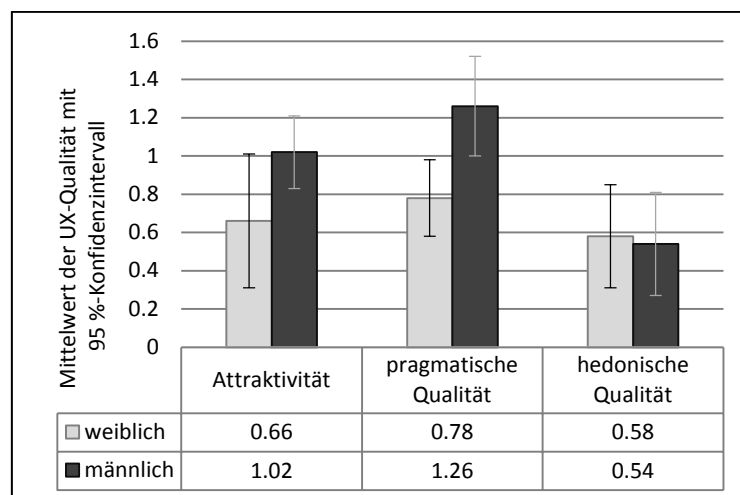


Abbildung 5-29: Mittelwerte der UX-Qualitäten Attraktivität, PQ und HQ bezogen auf das Geschlecht der Versuchsteilnehmer

5.2.2.2 Ergebnisse bezogen auf die psychische Beanspruchung

Für die Auswertung der psychischen Beanspruchung standen $N = 50$ Versuchsteilnehmer zur Verfügung. Zwei Teilnehmer konnten nicht herangezogen werden, da sie den Fragebogen nicht korrekt ausgefüllt hatten. Die Daten des Items 1, des Items 2 und des Items 7 waren normalverteilt und die Varianzen waren homogen (Anhang G2).

Die Abbildung 5-30 zeigt die Mittelwerte und die 95 %-Konfidenzintervalle aller Items des Fragebogens zur Beurteilung des Trainings, die mit einer Ratingskala erfasst wurden. Die Ergebnisse der Freitextitems 8 und 9 wurden bereits in Kapitel 5.2.1.2 dargestellt.

Der Vergleich der Gruppen zeigte Unterschiede in den erfassten Zieldimensionen *Beanspruchung* und *Attraktivität*. Die weiblichen Versuchsteilnehmer bewerteten die Anstrengungen und die Anforderungen, die für das Erlernen der Simulation notwendig waren (Item 1–3, Item 7), größer bzw. höher als die männlichen Versuchsteilnehmer. Die männlichen Versuchsteilnehmer hingegen bewerteten die farbliche Gestaltung und den Kontrast der Bedienoberflächen (Item 4 und Item 5) als angenehmer. Zudem empfanden die männlichen Versuchsteilnehmer die unterstützende Wirkung der farblichen Gestaltung bei der Fehlererkennung (Item 6) besser als die weiblichen Versuchsteilnehmer.

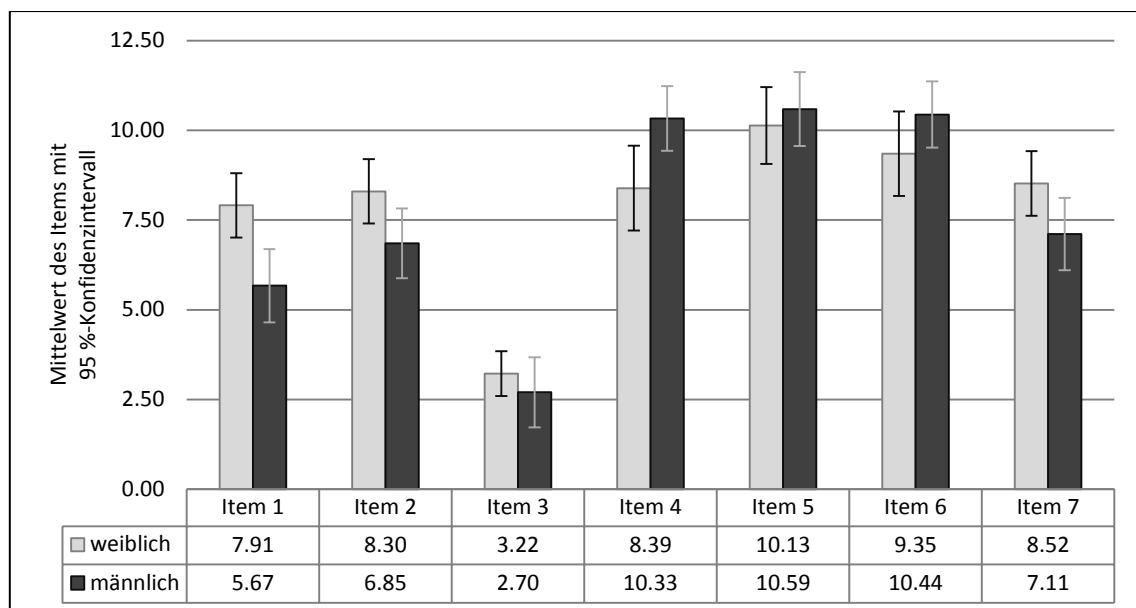


Abbildung 5-30: Mittelwerte und 95 %-Konfidenzintervalle der sieben Items bezogen auf das Geschlecht der Versuchsteilnehmer

Des Weiteren zeigt die Abbildung 5-30, dass die Konfidenzintervalle des Items 1 und des Items 4 nicht überlappten. Dies ließ auf einen signifikanten Unterschied schließen. Um dies zu überprüfen und um mögliche weitere signifikante Unterschiede feststellen zu können, wurden die Ergebnisse der Items im Einzelnen verglichen. Dabei werden die getrennt erfassten Dimensionen *Beanspruchung* und *Attraktivität* gesondert beschrieben.

5.2.2.2.1 Zieldimension *Beanspruchung*

Die Beanspruchung wurde mit den Items 1, 2, 3, 7 und dem Item 8 erfasst. Bei den ersten vier Items handelte es sich um Items, welche mit einer Ratingskala erfasst wurden. Item 8 war ein Freitextitem und wurde gesondert erfasst.

Item 1

Die weiblichen Versuchsteilnehmer empfanden die Anstrengung, die für das Erlernen der Simulation benötigt wurde, größer [weiblich ($M = 7.91$, $SD = 2.07$); männlich ($M = 5.67$, $SD = 2.59$)]. Dieser Unterschied stellte sich, sowohl mit dem t-Test ($t(48) = 3.351$, $p = .001$) als auch mit dem Mann-Whitney-U-Test ($U(27, 23) = 155$, $p = .002$), als signifikant heraus.

Item 2

Die geistigen Anforderungen wurden von den weiblichen Versuchsteilnehmern höher empfunden als von den männlichen Versuchsteilnehmern [weiblich ($M = 8.30$, $SD = 2.08$); männlich ($M = 6.85$, $SD = 2.46$)]. Dieser Unterschied stellte sich mit dem t-Test ($t(48) = 2.233$, $p = .030$) als signifikant und mit dem Mann-Whitney-U-Test ($U(27, 23) = 211.5$, $p = .052$) als marginal signifikant heraus. Da alle Voraussetzungen für die Nutzung des t-Tests erfüllt waren, wurde das Ergebnis des t-Tests akzeptiert. Demnach konnte ein signifikanter Unterschied festgestellt werden.

Item 3

Die körperlichen Anforderungen wurden von den weiblichen Versuchsteilnehmern höher eingeschätzt als von den männlichen Versuchsteilnehmern [weiblich ($M = 3.22$, $SD = 1.45$); männlich ($M = 2.70$, $SD = 2.48$)]. Der Unterschied konnte mit dem t-Test als nicht signifikant ($t(48) = 0.875$, $p = .386$) und mit dem Mann-Whitney-U-Test als signifikant ($U(27, 23) = 207$, $p = .039$) nachgewiesen werden. Da nur die Bedingung der Normalverteilung verletzt war und der t-Test gegenüber dieser Verletzung robust ist (Kapitel 5.1.6), wurde das Ergebnis des t-Tests stärker gewichtet. Demnach konnte kein signifikanter Unterschied festgestellt werden.

Item 7

Die Anstrengung, die notwendig war um die Bedienung zu erlernen, wurde von den weiblichen Versuchsteilnehmern größer eingeschätzt als von den männlichen Versuchsteilnehmern [weiblich ($M = 8.52$, $SD = 2.09$); männlich ($M = 7.11$, $SD = 2.56$)]. Dieser Unterschied stellte sich, sowohl mit dem t-Test ($t(48) = 2.110$, $p = .040$) als auch mit dem Mann-Whitney-U-Test ($U(27, 23) = 220.5$, $p = .076$), als signifikant heraus.

Item 8

Da nicht alle Versuchsteilnehmer die Möglichkeit in Anspruch nahmen das Freitextitem 8 für die Bewertung der Beanspruchung zu nutzen, stand nur eine begrenzte Anzahl an Antworten zur Verfügung. Wie bereits in Kapitel 5.2.1.2.2 erwähnt, wurde deshalb nur ein qualitativer Vergleich bezogen auf den Einfluss der farblichen Gestaltung durchgeführt. Der Versuchsleiter erwartete hier keine, durch das Geschlecht verursachte, Effekte. Aus diesem Grund wurde auf eine weitere Auswertung an dieser Stelle verzichtet.

5.2.2.2 Zieldimension *Attraktivität*

Die Attraktivität wurde mit den Items 4, 5, 6 und dem Item 9 erfasst. Bei den ersten drei Items handelte es sich um Items, welche mit einer Ratingskala erfasst wurden. Item 9 war ein Freitextitem und wurde gesondert erfasst.

Item 4

Den weiblichen Versuchsteilnehmern gefiel die farbliche Gestaltung der Bedienoberflächen insgesamt weniger als den männlichen Versuchsteilnehmern [weiblich ($M = 8.39$, $SD = 2.74$); männlich ($M = 10.33$, $SD = 2.29$)]. Dieser Unterschied stellte sich, sowohl mit dem t-Test ($t(48) = -2.731$, $p = .009$) als auch mit dem U-Test ($U(27, 23) = 176.5$, $p = .008$), als signifikant heraus.

Item 5

Der Kontrast der Bedienoberflächen gefiel den weiblichen Versuchsteilnehmern insgesamt weniger als den männlichen Versuchsteilnehmern [weiblich ($M = 10.13$, $SD = 2.47$); männlich ($M = 10.59$, $SD = 2.61$)]. Dieser Unterschied stellte sich als nicht signifikant heraus (Anhang G2).

Item 6

Die unterstützende Wirkung der farblichen Gestaltung der Bedienoberflächen bei der Fehlererkennung wurde von den weiblichen Versuchsteilnehmern geringer eingeschätzt als von den männlichen Versuchsteilnehmern [weiblich ($M = 9.35$, $SD = 2.72$); männlich ($M = 10.44$, $SD = 2.34$)]. Dieser Unterschied stellte sich als nicht signifikant heraus (Anhang G2).

Item 9

Da nicht alle Versuchsteilnehmer die Möglichkeit in Anspruch nahmen das Freitextitem 9 für die Bewertung der Beanspruchung zu nutzen, stand nur eine begrenzte Anzahl an Antworten zu Verfügung. Wie bereits in Kapitel 5.2.1.2.2 erwähnt, wurde deshalb nur ein qualitativer Vergleich bezogen auf den Einfluss der farblichen Gestaltung durchgeführt. Der Versuchslei-

ter erwartete hier keine, durch das Geschlecht verursachte, Effekte. Aus diesem Grund wurde auf eine weitere Auswertung an dieser Stelle verzichtet.

5.2.2.3 Ergebnisse bezogen auf die Befindlichkeit

Als Datengrundlage für die Auswertung der Befindlichkeit standen $N = 52$ Versuchsteilnehmer zur Verfügung. Die Daten waren nicht normalverteilt, aber die Varianzen waren homogen (Anhang G3).

Analog zu der Auswertung bezogen auf den Einfluss der farblichen Gestaltung (Kapitel 5.2.1.3), wurden zu Beginn die Differenzen der Vorher-Nachher-Messung der verschiedenen Dimensionen gebildet und als vorläufiges Maß für den Einfluss des Geschlechts der Versuchsteilnehmer auf die Befindlichkeit genutzt (Abbildung 5-31).

Der Vergleich der Differenzen zeigte, dass die Wirkrichtung der Dimensionen unabhängig vom Geschlecht war. Sowohl bei den weiblichen als auch bei den männlichen Versuchsteilnehmern stellte sich eine Verschlechterung in den Dimensionen *GS* und *RU* ein sowie eine Verbesserung der Befindlichkeit in der Dimension *WM*. Des Weiteren konnten zwei signifikante Effekte in der Dimension *RU* nachgewiesen werden. Der erste signifikante Unterschied betraf die Nachher-Messung der Dimension *RU* und der zweite die Differenz der Vorher-Nachher-Messung in der Gruppe mit den weiblichen Versuchsteilnehmern.

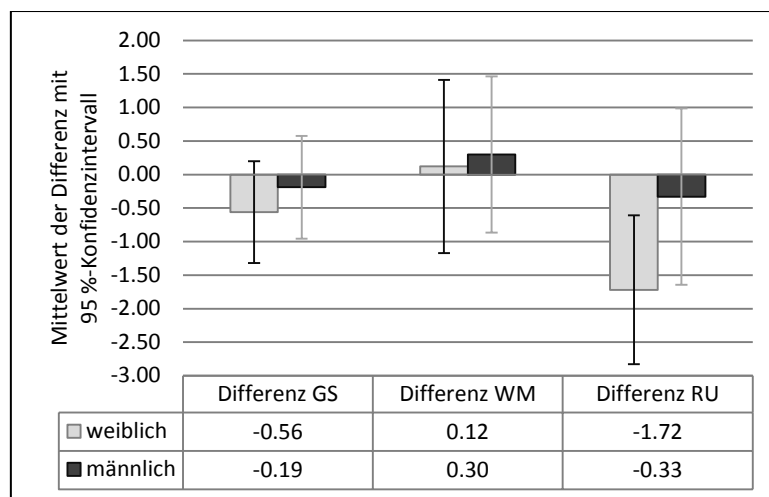


Abbildung 5-31: Mittelwerte und 95 %-Konfidenzintervalle der Differenzen der Dimensionen *Gute-Schlechte-Stimmung* (GS), *Wachheit-Müdigkeit* (WM) und *Ruhe-Unruhe* (RU) bezogen auf das Geschlecht der Versuchsteilnehmer

5.2.2.3.1 Zieldimension *Gute-Schlechte-Stimmung* (GS)

Abbildung 5-32 verdeutlicht die Gesamtveränderung der Dimension *GS*. Die weiblichen Versuchsteilnehmer bewerteten die Dimension *GS* grundsätzlich schlechter als die männlichen Versuchsteilnehmer. Diese Beurteilung war ausschließlich qualitativer Natur, d. h. es konnte kein signifikanter Unterschied festgestellt werden (Anhang G3).

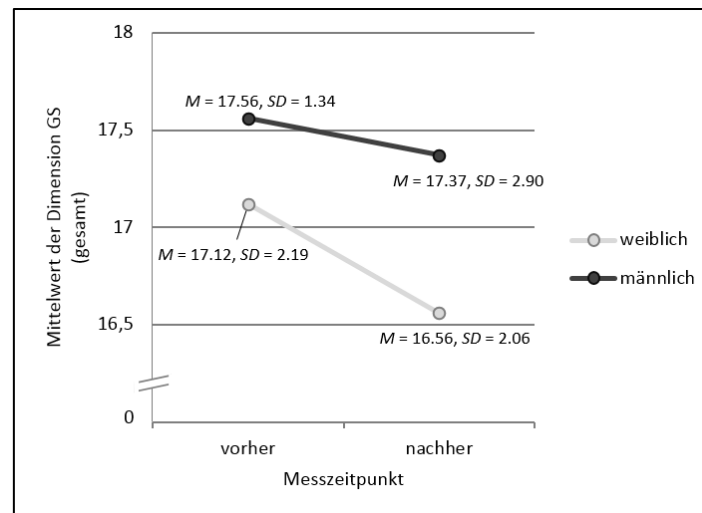


Abbildung 5-32: Mittelwerte und Standardabweichungen der Vorher-Nachher-Messung der Dimension *Gute-Schlechte-Stimmung* (GS) bezogen auf das Geschlecht der Versuchsteilnehmer

Auch wenn kein signifikanter Unterschied nachweisbar war, wurde der Ursprung für die qualitative Abweichung untersucht. Dafür wurde die Gesamtveränderung der Dimension GS wieder in ihren positiven und ihren negativen Aspekt zerlegt. Abbildung 5-33 zeigt das Ergebnis.

Der Vergleich des positiven Aspekts ergab bei den weiblichen Versuchsteilnehmern keine Veränderung. Sie bewerteten die Dimension GS nach der Eingewöhnungsphase schlechter als die männlichen Versuchsteilnehmer. Bei den männlichen Versuchsteilnehmern ergab sich insgesamt keine Änderung in der Vorher-Nachher-Messung für den positiven Aspekt.

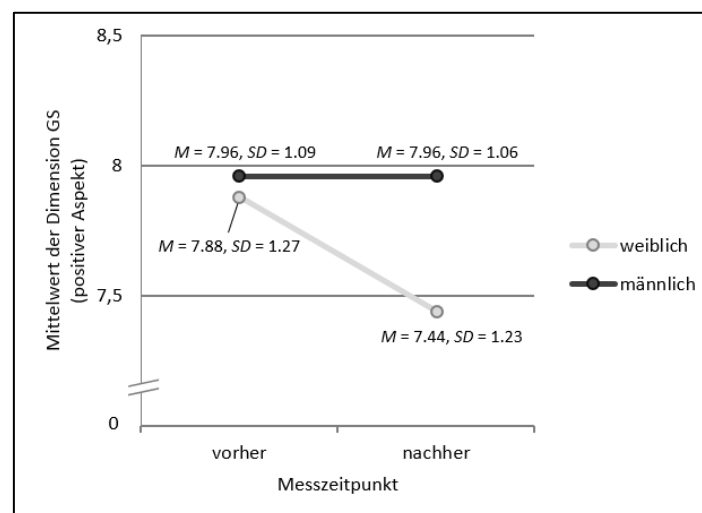


Abbildung 5-33: Mittelwerte und Standardabweichungen der Vorher-Nachher-Messung des positiven Aspekts der Dimension *Gute-Schlechte-Stimmung* (GS) bezogen auf das Geschlecht der Versuchsteilnehmer

Als nächstes wurde der negative Aspekt der Dimension GS untersucht (Abbildung 5-34). Sowohl die weiblichen als auch die männlichen Versuchsteilnehmer bewerteten den negativen Aspekt schlechter.

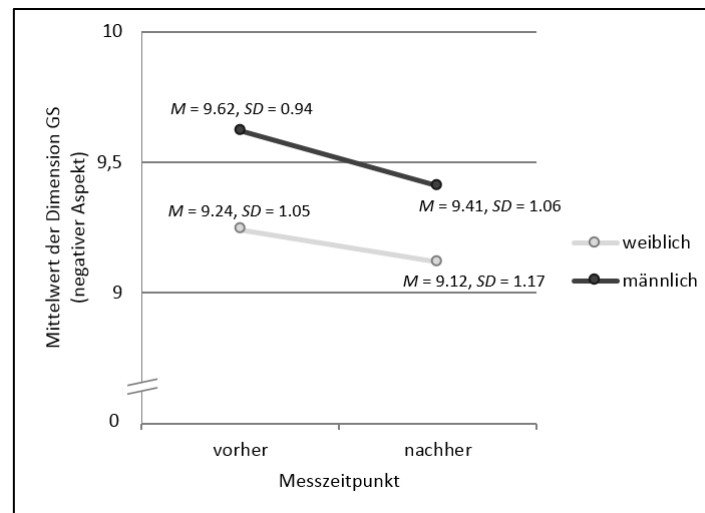


Abbildung 5-34: Mittelwerte und Standardabweichungen der Vorher-Nachher-Messung des negativen Aspekts der Dimension *Gute-Schlechte-Stimmung* (GS) bezogen auf das Geschlecht der Versuchsteilnehmer

Zusammenfassend bewerteten beide Geschlechter die Befindlichkeit in der Dimension GS nach der Eingewöhnungsphase schlechter. Die Ergebnisse des positiven Aspekts wurden durch die männlichen Versuchsteilnehmer nicht beeinflusst, weshalb die Gesamtveränderung in diesem Fall etwas geringer ausfiel.

5.2.2.3.2 Zieldimension *Wachheit-Müdigkeit* (WM)

Der Vergleich der Ergebnisse der Dimension WM zeigte eine Verbesserung der Befindlichkeit. Bei beiden Geschlechtern war nach der Eingewöhnungsphase eine positivere Stimmung messbar. Sie waren wacher. Abbildung 5-35 verdeutlicht die Gesamtveränderung in der Dimension WM. Wie auch bei der Dimension GS war der Unterschied jedoch nicht signifikant (Anhang G3).

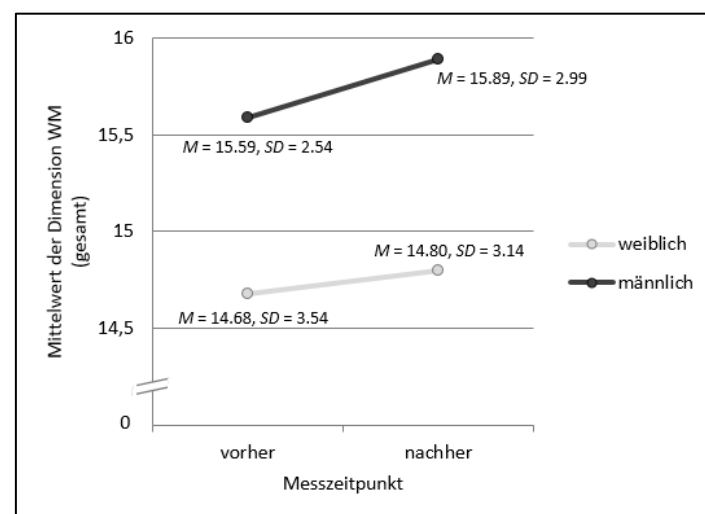


Abbildung 5-35: Mittelwerte und Standardabweichungen der Vorher-Nachher-Messung der Dimension *Wachheit-Müdigkeit* (WM) bezogen auf das Geschlecht der Versuchsteilnehmer

Der Vergleich des positiven Aspekts ergab einen Unterschied in der Wirkrichtung zwischen den Geschlechtern (Abbildung 5-36). Die männlichen Versuchsteilnehmer bewerteten den positiven Aspekt der Dimension *WM* nach der Eingewöhnungsphase besser. Bei den weiblichen Versuchsteilnehmern zeigte sich ein anderes Bild. Sie bewerteten den positiven Aspekt nach der Eingewöhnungsphase schlechter.

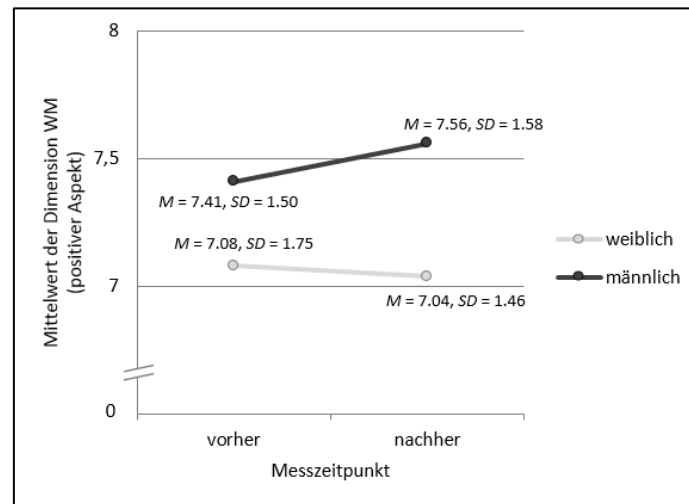


Abbildung 5-36: Mittelwerte und Standardabweichungen der Vorher-Nachher-Messung des positiven Aspekts der Dimension *Wachheit-Müdigkeit* (WM) bezogen auf das Geschlecht der Versuchsteilnehmer

Als nächstes wurde der negative Aspekt der Dimension *WM* untersucht (Abbildung 5-37). Der Vergleich zeigte einen positiven Einfluss des Geschlechts der Versuchsteilnehmer. Sowohl die weiblichen als auch die männlichen Versuchsteilnehmer bewerteten den negativen Aspekt der Dimension *WM* nach der Eingewöhnungsphase besser. Die Befindlichkeit hatte sich verbessert und die Teilnehmer waren nach der Eingewöhnungsphase weniger müde.

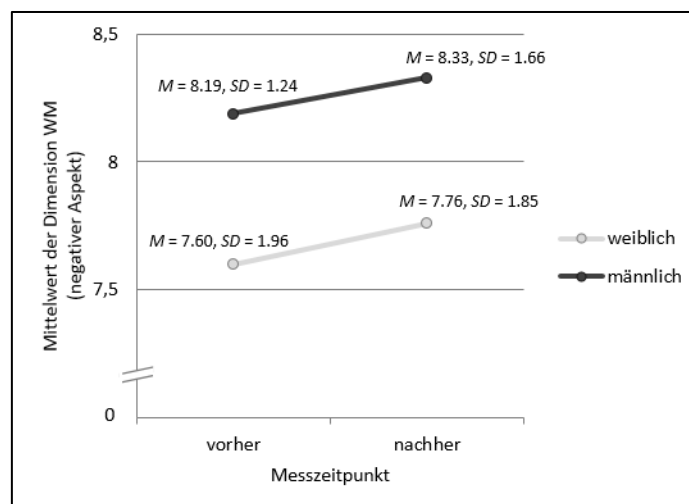


Abbildung 5-37: Mittelwerte und Standardabweichungen der Vorher-Nachher-Messung des negativen Aspekts der Dimension *Wachheit-Müdigkeit* (WM) bezogen auf das Geschlecht der Versuchsteilnehmer

Zusammenfassend bewerteten beide Geschlechter die Dimension *WM* nach der Eingewöhnungsphase besser als zu Beginn des Versuchs. Auffällig war, dass die weiblichen Versuchs-

teilnehmer den positiven Aspekt negativ bewerteten. Dadurch fiel die Gesamtdifferenz geringer aus als bei den männlichen Versuchsteilnehmern.

5.2.2.3.3 Zieldimension *Ruhe-Unruhe* (RU)

In der Dimension *RU* machte sich der Einfluss des Geschlechts der Versuchsteilnehmer stärker bemerkbar. Die weiblichen Versuchsteilnehmer bewerteten die Dimension *RU* nach der Eingewöhnungsphase sehr viel schlechter als die männlichen Versuchsteilnehmer (Abbildung 5-38).

Im Gegensatz zu den Dimensionen *GS* und *WM* ergaben sich in der Dimension *RU* zwei signifikante Unterschiede. Der erste Unterschied betraf die Nachher-Messung zwischen den Geschlechtern. Dieser Unterschied stellte sich, sowohl mit dem t-Test ($t(50) = -2.600, p = .012$) als auch mit dem Mann-Whitney-U-Test ($U(27, 25) = 197.5, p = .010$), als signifikant heraus. Der zweite Effekt war ein signifikanter Unterschied in der Vorher-Nachher-Messung der Gruppe mit den weiblichen Versuchsteilnehmern. Dieser konnte sowohl mit dem t-Test für abhängige Stichproben ($t(24) = 3.203, p = .004$) als auch mit dem Wilcoxon-Test ($T(25) = 41.5, p = .003, z = -2.956$) nachgewiesen werden.⁸⁶ Um zu überprüfen, ob die beiden nachgewiesenen Unterschiede voneinander abhängig waren und/oder ob diese aus einer Wechselbeziehung zwischen farblicher Gestaltung und Geschlecht herrührten, wurde dies zusätzlich in Kapitel 5.2.3.3 inferenzstatistisch untersucht.

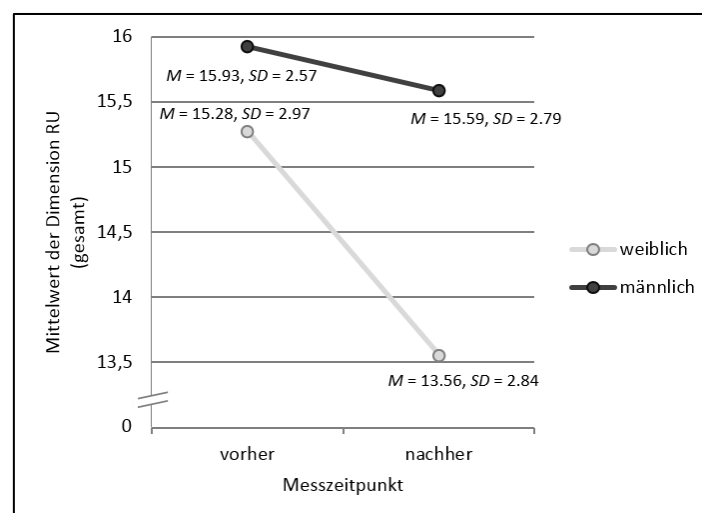


Abbildung 5-38: Mittelwerte und Standardabweichungen der Vorher-Nachher-Messung der Dimension *Ruhe-Unruhe* (RU) bezogen auf das Geschlecht der Versuchsteilnehmer

Im nächsten Schritt wurde die Gesamtveränderung in den positiven und den negativen Aspekt der Dimension zerlegt (Abbildung 5-39). Für den positiven Aspekt ergab sich ein interessantes Bild. Es zeigte sich ein Unterschied in der Wirkrichtung zwischen den Geschlechtern.

⁸⁶ Um einen möglichen Effekt zwischen den Ergebnissen der Vorher-Nachher-Messung zu untersuchen, wurde ein Test für Messwiederholung verwendet. Als parametrisches Verfahren wurde der t-Test für abhängige Stichproben verwendet und als nicht parametrisches Verfahren der Wilcoxon-Test.

Die männlichen Versuchsteilnehmer bewerteten den positiven Aspekt nach der Eingewöhnungsphase besser und die weiblichen Versuchsteilnehmer schlechter.

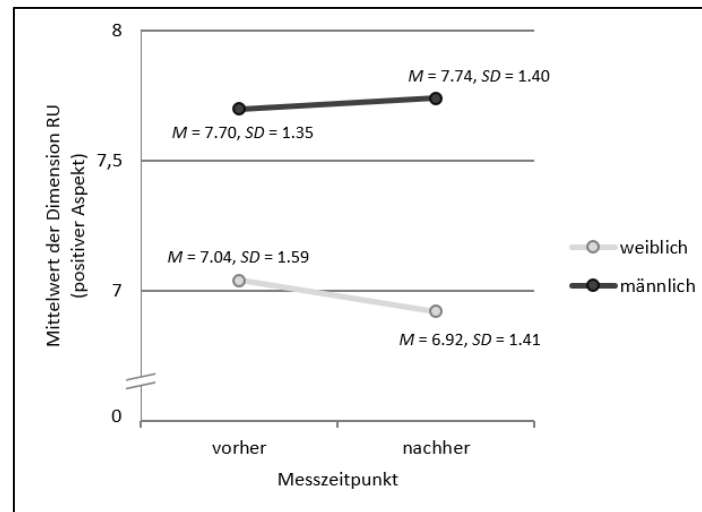


Abbildung 5-39: Mittelwerte und Standardabweichungen der Vorher-Nachher-Messung des positiven Aspekts der Dimension *Ruhe-Unruhe* (RU) bezogen auf das Geschlecht der Versuchsteilnehmer

Der Vergleich des negativen Aspekts der Dimension *RU* zeigte ein ähnliches Bild (Abbildung 5-40). Die weiblichen Versuchsteilnehmer bewerteten diesen Aspekt sehr viel schlechter als die männlichen Versuchsteilnehmer.

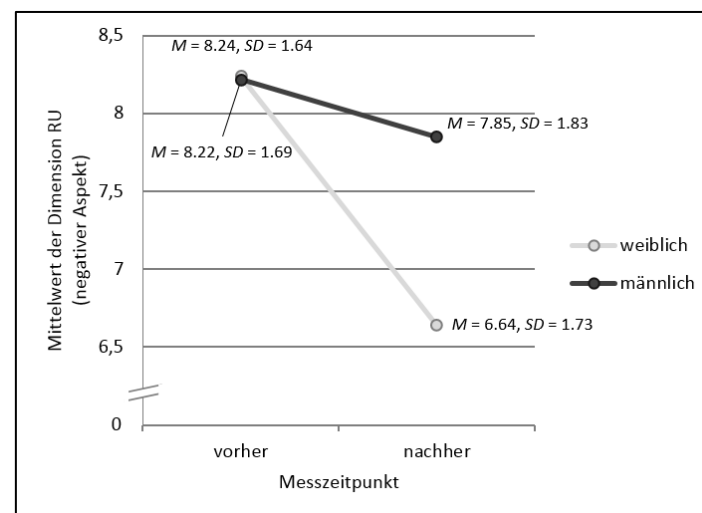


Abbildung 5-40: Mittelwerte und Standardabweichungen der Vorher-Nachher-Messung des negativen Aspekts der Dimension *Ruhe-Unruhe* (RU) bezogen auf das Geschlecht der Versuchsteilnehmer

Zusammenfassend bewerteten die weiblichen Versuchsteilnehmer die Ergebnisse der Dimension *RU* wesentlich schlechter als die männlichen Versuchsteilnehmer. Vor allem der negative Aspekt wurde von den weiblichen Versuchsteilnehmern sehr negativ bewertet. Um zu überprüfen, ob diese Unterschiede aus einer Wechselbeziehung zwischen farblicher Gestaltung und Geschlecht herrührten, wurde dies zusätzlich in Kapitel 5.2.3.3 inferenzstatistisch untersucht (Kapitel 5.2.3.3).

5.2.2.4 Ergebnisse bezogen auf die Bedienungssicherheit

Als Datengrundlage wurden $N = 50$ Versuchsteilnehmer herangezogen. Zwei Teilnehmer konnten für die Auswertung nicht verwendet werden, da sie nicht für jeden Messpunkt eine Bewertung abgegeben hatten. Aus diesem Grund wurden sie bei der Betrachtung dieser abhängigen Variable komplett ausgeschlossen, um für alle Messpunkte die gleiche Anzahl an Antworten zu haben. Die Daten waren nicht normalverteilt, aber die Varianzen waren homogen (Anhang G4).

Der Vergleich des Verlaufs der Bedienungssicherheit zeigte, dass sich die männlichen Versuchsteilnehmer grundsätzlich sicherer fühlten als die weiblichen Versuchsteilnehmer (Abbildung 5-41).

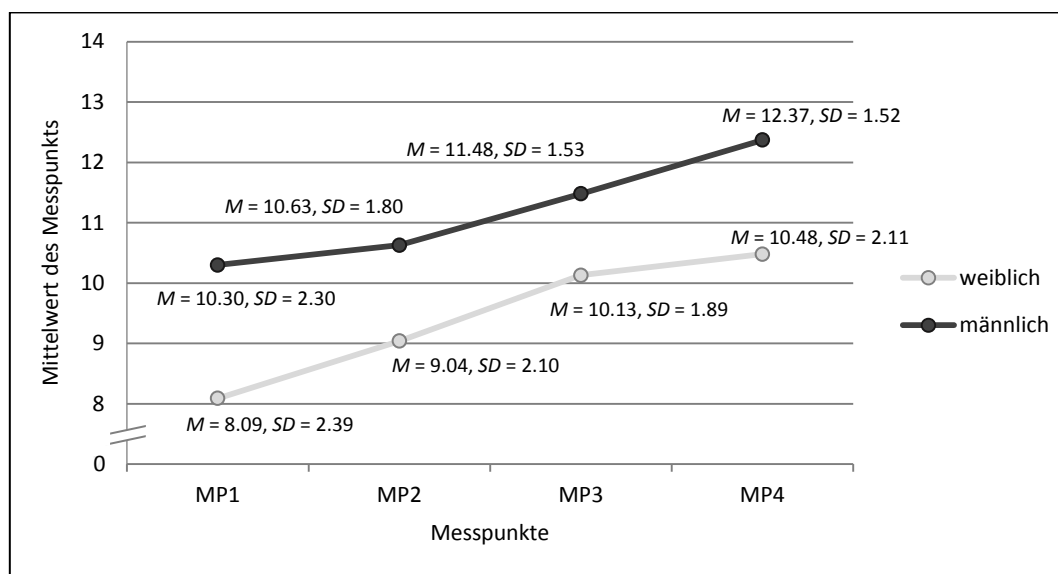


Abbildung 5-41: Mittelwerte und Standardabweichungen der vier Messpunkte für die Bedienungssicherheit bezogen auf das Geschlecht der Versuchsteilnehmer

Gerade zu Beginn und am Ende der Eingewöhnungsphase zeigte sich ein großer Unterschied bei der Selbsteinschätzung. Unabhängig davon, konnten die Unterschiede für alle Messpunkte sowohl mit dem t-Test als auch mit dem Mann-Whitney-U-Test als signifikant nachgewiesen werden (Tabelle 5-12). Das Geschlecht der Versuchsteilnehmer hatte somit einen signifikanten Einfluss auf die Beurteilung der Bedienungssicherheit.

Tabelle 5-12: Interferenzstatistische Ergebnisse für die vier Messpunkte

Messpunkt	t-Test	U-Test
MP1	$t(48) = -3.323, p = .002$	$U(27, 23) = 160, p = .003; z = -2.950$
MP2	$t(48) = -2.877, p = .006$	$U(27, 23) = 165, p = .004; z = -2.881$
MP3	$t(48) = -2.794, p = .007$	$U(27, 23) = 184.5, p = .012; z = -2.499$
MP4	$t(48) = -3.675, p < .001$	$U(27, 23) = 148.5, p < .001; z = -3.193$

Obwohl sich die weiblichen Versuchsteilnehmer in jedem Messpunkt schlechter einschätzten als die männlichen, zeigte die Differenzbetrachtung in Abbildung 5-42, dass der Sicherheitszuwachs bei den weiblichen Versuchsteilnehmern in fast allen Messpunkten größer war. Dies führte dazu, dass der Zuwachs bei den weiblichen Versuchsteilnehmern insgesamt größer war – was die Differenz zwischen dem ersten und dem letzten Messpunkt (MP4/MP1) zusätzlich bestätigte.

Zudem zeigte sich eine weitere Auffälligkeit. Der Sicherheitszuwachs lag bei den weiblichen Versuchsteilnehmern eher am Anfang des Experiments und bei den männlichen Versuchsteilnehmern eher am Ende. Gerade in der Eingewöhnungsphase zeigte sich bei den weiblichen Versuchsteilnehmern ein Zuwachs (MP2/MP1 und MP3/MP2). Bei den männlichen Versuchsteilnehmern führte der letzte Trainingsblock (MP3/MP2) und der standardisierte Testlauf (MP4/MP3) zu einem größeren Sicherheitszuwachs. Die gefundenen Unterschiede bei den Differenzen waren ausschließlich qualitativer Natur (Anhang G4).

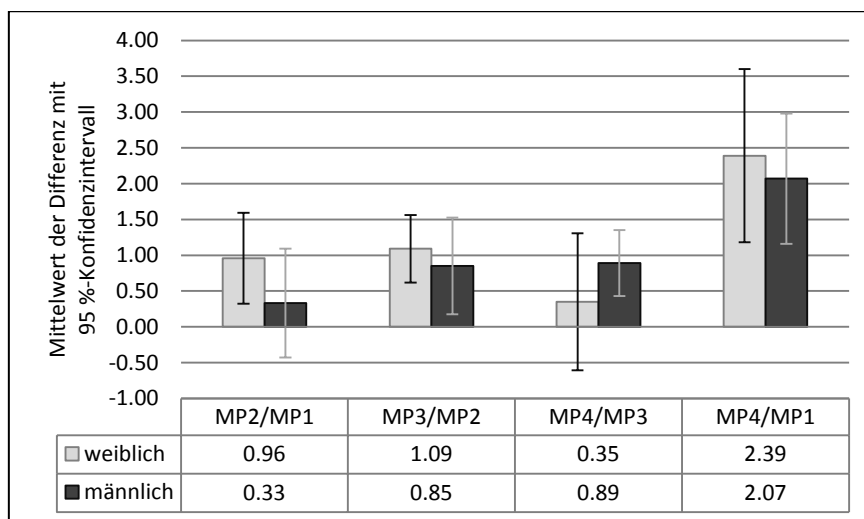


Abbildung 5-42: Mittelwerte und 95 %-Konfidenzintervalle der Messpunktdifferenzen bezogen auf das Geschlecht der Versuchsteilnehmer

5.2.2.5 Ergebnisse bezogen auf die vordefinierten Leistungskennzahlen

Wie in Kapitel 5.2.1.5 bereits beschrieben, standen für die Auswertung der Leistungskennzahlen 12 Versuchsteilnehmer nicht zur Verfügung. Die 17 weiblichen Versuchsteilnehmer waren zwischen 19 und 31 Jahre alt und die 23 männlichen Versuchsteilnehmer zwischen 19 und 32 Jahren. Bei den weiblichen Versuchsteilnehmern betrug das Durchschnittsalter 22.94 Jahre ($SD = 3.07$) und bei den männlichen 24.61 Jahre ($SD = 3.62$).

5.2.2.5.1 Anzahl der abgefertigten Zugfahrten

Die erfassten Daten für die Untersuchung der abgefertigten Zugfahrten waren nicht normalverteilt, aber die Varianzen waren homogen (Anhang G5).

Die weiblichen Versuchsteilnehmer fertigten im Mittel weniger Zugfahrten ab als die männlichen Versuchsteilnehmer [weiblich ($M = 44.47$, $SD = 7.72$); männlich ($M = 48.22$, $SD = 2.94$)]. Der Unterschied konnte nur mit dem t-Test ($t(19.45) = -1.903$, $p = .072$) als marginal signifikant nachgewiesen werden. Da zwei der drei Voraussetzungen für die Verwendung des t-Tests nicht erfüllt waren, wurde das Ergebnis des Mann-Whitney-U-Tests ($U(23, 17) = 160$, $p = .342$) stärker gewichtet. Demnach konnte kein signifikanter Unterschied nachgewiesen werden.

5.2.2.5.2 Fehlerentdeckungszeit

Zuerst wurden die Gesamtfehlerentdeckungszeiten miteinander verglichen. Hierfür wurden die Einzelzeiten aufsummiert. Die Daten waren nicht normalverteilt und die Varianzen waren nicht homogen (Anhang G5).

Die weiblichen Versuchsteilnehmer benötigten im Mittel länger, um die Unregelmäßigkeiten zu erkennen [weiblich ($M = 648.97$, $SD = 217.62$); männlich ($M = 632.91$, $SD = 297.94$)]. Dieser Unterschied konnte nicht als signifikant nachgewiesen werden (Anhang G5).

Wie auch bei der Untersuchung des Einflusses der farblichen Gestaltung ergab die Analyse der Einzelereignisse, bis auf ein paar Auffälligkeiten, keine aussagekräftigen Unterschiede. Abbildung 5-43 zeigt die Mittelwerte der Fehlerentdeckungszeiten der einzelnen Unregelmäßigkeiten grafisch und Tabelle 5-13 gibt einen Überblick über die Zahlenwerte. Im Folgenden werden nur die Unregelmäßigkeiten weiter untersucht, bei denen der Unterschied zwischen beiden Gruppen am aussagekräftigsten war.

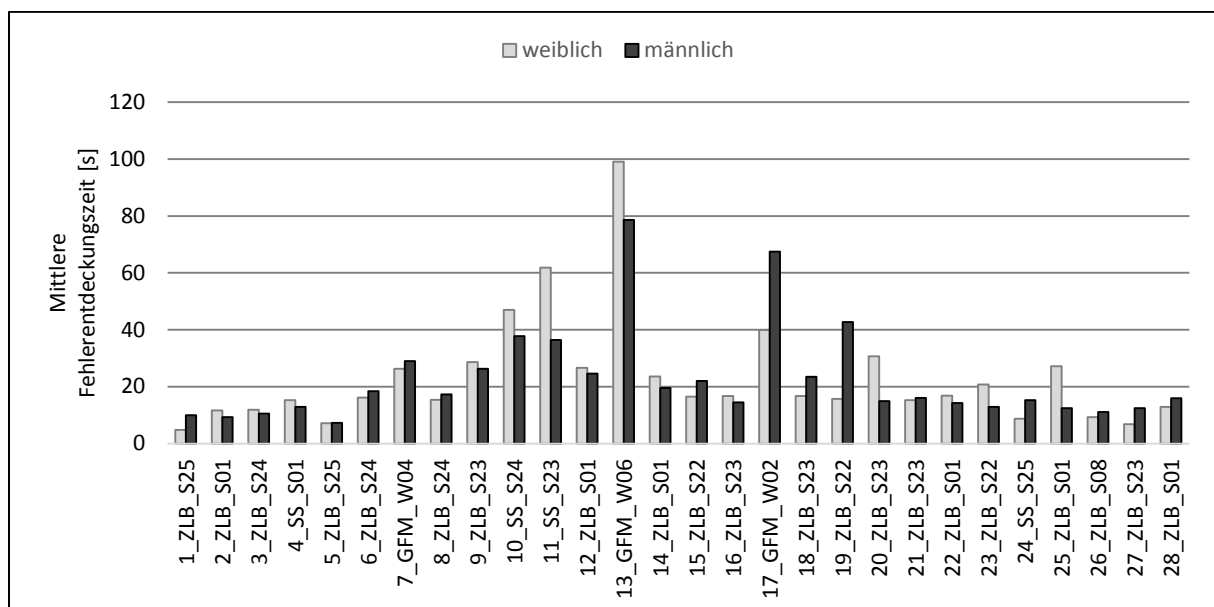


Abbildung 5-43: Mittelwerte der Fehlerentdeckungszeiten für die einzelnen Unregelmäßigkeiten bezogen auf das Geschlecht der Versuchsteilnehmer

Wie auch bei den Ergebnissen in Kapitel 5.2.1.5.2, wurde ein Unterschied bei den Fehlerentdeckungszeiten der Signalstörungen festgestellt. Bei drei von vier Signalstörungen benötigten die weiblichen Versuchsteilnehmer länger für das Erkennen. Gerade die Unregelmäßigkeit *11_SS_S23* war dabei sehr auffällig. Mit Blick auf die Versuchsgruppen fiel auf, dass in der Gruppe mit der dunklen Bedienoberfläche die Anzahl der weiblichen Versuchsteilnehmer im Vergleich zu den männlichen größer war. Um zu überprüfen, ob dieser Umstand einen Einfluss hatte, wurde dies zusätzlich in Kapitel 5.2.3.5 untersucht.

Bei Unregelmäßigkeit *17_GFM_W02* zeigte sich auch ein Unterschied zwischen den Geschlechtern. Die männlichen Versuchsteilnehmer benötigten im Mittel länger, um die Gleisfreimelde störung zu erkennen. Zudem wurde die Unregelmäßigkeit von den Versuchsteilnehmern insgesamt fünfmal nicht erkannt – viermal davon von den männlichen Versuchsteilnehmern. Da in der Gruppe mit der hellen Bedienoberfläche der Anteil der männlichen Versuchsteilnehmer größer war, wurde in Kapitel 5.2.3.5 überprüft, ob dieser Umstand einen Einfluss hatte.

Der Ursprung für die Unterschiede bei der mittleren Fehlerentdeckungszeit der Unregelmäßigkeiten *18_ZLB_S23*, *19_ZLB_S22*, *20_ZLB_S23* und *25_ZLB_S01* konnte nicht nachvollzogen werden. Beide Gruppen waren in gleichem Maße abgelenkt und die Unregelmäßigkeit wurde von allen Versuchsteilnehmern erkannt. Um einen Zusammenhang zwischen der farblichen Gestaltung und dem Geschlecht der Versuchsteilnehmer ausschließen zu können, wurden zusätzlich die Wechselwirkungen zwischen beiden unabhängigen Variablen in Kapitel 5.2.3.5 untersucht.

Tabelle 5-13: Mittelwerte und Standardabweichungen der Fehlerentdeckungszeiten für die einzelnen Unregelmäßigkeiten bezogen auf das Geschlecht der Versuchsteilnehmer

Nr.	weiblich (N = 17)		männlich (N = 23)		Nr.	weiblich (N = 17)		männlich (N = 23)	
	M	SD	M	SD		M	SD	M	SD
<i>1_ZLB_S25</i>	4.74	4.88	9.91	13.46	<i>15_ZLB_S22</i>	16.44	19.32	21.96	25.79
<i>2_ZLB_S01</i>	11.65	18.66	9.30	9.13	<i>16_ZLB_S23</i>	16.77	23.01	14.41	15.13
<i>3_ZLB_S24</i>	11.82	8.38	10.48	8.21	<i>17_GFM_W02</i>	39.94	56.63	67.44	68.81
<i>4_SS_S01</i>	15.29	16.19	12.87	18.67	<i>18_ZLB_S23</i>	16.74	15.63	23.50	31.11
<i>5_ZLB_S25</i>	7.15	6.74	7.22	7.27	<i>19_ZLB_S22</i>	15.68	15.94	42.67	54.55
<i>6_ZLB_S24</i>	16.12	17.38	18.41	26.73	<i>20_ZLB_S23</i>	30.65	51.40	14.89	15.20
<i>7_GFM_W04</i>	26.29	31.67	29.02	37.65	<i>21_ZLB_S23</i>	15.24	18.71	16.09	19.66
<i>8_ZLB_S24</i>	15.41	19.88	17.33	17.20	<i>22_ZLB_S01</i>	16.88	19.98	14.20	16.06
<i>9_ZLB_S23</i>	28.59	35.84	26.30	24.69	<i>23_ZLB_S22</i>	20.74	22.58	12.87	14.93
<i>10_SS_S24</i>	46.94	50.00	37.78	33.96	<i>24_SS_S25</i>	8.74	10.40	15.20	19.39
<i>11_SS_S23</i>	61.82	58.33	36.44	45.62	<i>25_ZLB_S01</i>	27.21	28.94	12.46	9.00
<i>12_ZLB_S01</i>	26.59	28.49	24.57	23.19	<i>26_ZLB_S08</i>	9.24	10.31	11.07	10.40
<i>13_GFM_W06</i>	99.15	67.57	78.59	64.11	<i>27_ZLB_S23</i>	6.77	12.95	12.46	13.39
<i>14_ZLB_S01</i>	23.53	24.09	19.54	16.79	<i>28_ZLB_S01</i>	12.88	8.80	15.96	17.47

Des Weiteren wurde bei der Unregelmäßigkeit *7_GFM_W04* in Kapitel 5.2.1.5.2 ein Einfluss des Geschlechts vermutet. Wie Abbildung 5-43 zeigt, war die Unregelmäßigkeit bezogen auf das Geschlecht der Versuchsteilnehmer unauffällig. Demnach bestand kein Einfluss einer weiteren Variable und der nachgewiesene Unterschied zwischen den Gruppen wurde ausschließlich durch die farbliche Gestaltung der Bedienoberfläche hervorgerufen.

5.2.2.5.3 Verhältnis zwischen Klickzahl und gemachten Fehlern in der Nebenaufgabe

Durch technische Probleme konnte ein Versuchsteilnehmer für die Auswertung nicht herangezogen werden. Die Daten waren nicht normalverteilt und die Varianzen waren homogen (Anhang G5).

Die weiblichen Versuchsteilnehmer erzielten ein schlechteres Verhältnis zwischen der Klickzahl und den gemachten Fehlern als die männlichen Versuchsteilnehmer [weiblich ($M = 151.19$, $SD = 107.35$); männlich ($M = 144.91$, $SD = 124.74$)]. Der Unterschied stellte sich bezogen auf das Geschlecht der Versuchsteilnehmer als nicht signifikant heraus.

5.2.2.5.4 Anzahl der Hilfshandlungen

Für die Auswertung wurde die Gesamtanzahl der Hilfshandlungen verwendet. Die erfassten Daten waren nicht normalverteilt, aber die Varianzen waren homogen (Anhang G5).

Die weiblichen Versuchsteilnehmer führten weniger Hilfshandlungen aus als die männlichen Versuchsteilnehmer. [weiblich ($M = 9.88$, $SD = 2.76$); männlich ($M = 10.83$, $SD = 2.81$)]. Dieser Unterschied konnte nicht als signifikant nachgewiesen werden (Anhang G5). Deshalb erfolgte keine weitere Untersuchung der Einzelbestandteile *VS-Fahrt Weiche* und *VS-Fahrt Signal*.

5.2.2.5.5 Anzahl der Bedienfehler

Für die Auswertung wurde die Gesamtanzahl der Bedienfehler verwendet. Die erfassten Daten waren nicht normalverteilt, aber die Varianzen waren homogen (Anhang G5).

Während des Hauptversuchs machten die weiblichen Versuchsteilnehmer mehr Bedienfehler als die männlichen Versuchsteilnehmer [weiblich ($M = 6.53$, $SD = 4.75$); männlich ($M = 4.52$, $SD = 4.21$)]. Dieser Unterschied war nicht signifikant (Anhang G5). Aus diesem Grund erfolgte keine weitere Untersuchung bezogen auf die Einzelklassifizierung der Bedienfehler.

5.2.2.5.6 Gesamtüberwachungszeit

Die erfassten Daten waren nicht normalverteilt und die Varianzen waren nicht homogen (Anhang G5).

Die weiblichen Versuchsteilnehmer verbrachten mehr Zeit mit der Überwachung des Bedienbereichs als die männlichen Versuchsteilnehmer [weiblich ($M = 16.36$, $SD = 2.95$); männlich ($M = 15.83$, $SD = 2.98$)]. Dieser Unterschied war nicht signifikant (Anhang G5).

5.2.3 Wechselwirkung zwischen den unabhängigen Variablen

In den vorangegangenen Kapiteln wurden die Ergebnisse bezogen auf den Einfluss der farblichen Gestaltung und das Geschlecht der Versuchsteilnehmer dargestellt. Bei der Untersuchung dieser Einflüsse wurde festgestellt, dass einige Unterschiede sowohl durch die farbliche Gestaltung als auch durch das Geschlecht der Versuchsteilnehmer hervorgerufen wurden. Um Zusammenhänge zwischen den beiden unabhängigen Variablen ausschließen zu können, wurden mögliche Wechselwirkungen zwischen den unabhängigen Variablen untersucht.⁸⁷

5.2.3.1 Ergebnisse bezogen auf die User Experience

Bei der User Experience wurden in der Studie keine signifikanten Unterschiede festgestellt. Demnach konnten auch keine Zusammenhänge zwischen den unabhängigen Variablen nachgewiesen werden (Anhang G1). Nichtsdestotrotz wurde eine qualitative Untersuchung des Zusammenhangs der unabhängigen Variablen anhand von grafischen Auswertungen durchgeführt. Die Ergebnisse dieser Analyse werden in den folgenden Abbildungen dargestellt (Abbildung 5-44 bis Abbildung 5-49).

Die Untersuchung der Dimension *Attraktivität* ergab, dass die männlichen Versuchsteilnehmer die Attraktivität beider Bedienoberflächen grundsätzlich besser bewerteten als die weiblichen Versuchsteilnehmer (Abbildung 5-44). Es konnte eine Tendenz bezogen auf das Geschlecht der Versuchsteilnehmer nachgewiesen werden. Die weiblichen Versuchsteilnehmer präferierten die helle Bedienoberfläche und die männlichen Versuchsteilnehmer die dunkle Bedienoberfläche.

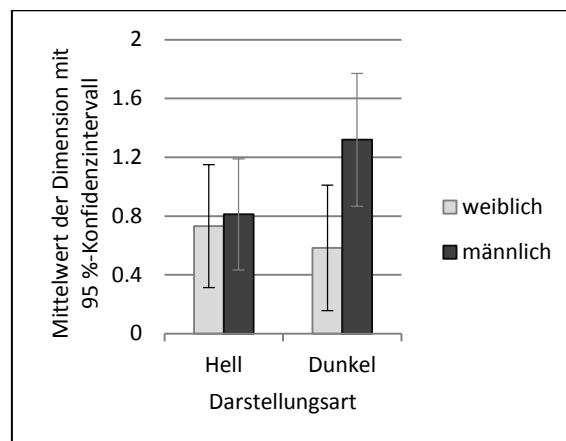


Abbildung 5-44: Mittelwerte und 95 %-Konfidenzintervalle der Dimension *Attraktivität* bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen

⁸⁷ Hierfür wurde eine mehrfaktorielle univariate/multivariate Varianzanalyse durchgeführt.

In der Dimension *Durchschaubarkeit* zeigte sich ein ähnliches Bild. Wie Abbildung 5-45 zeigt, bewerteten die männlichen Versuchsteilnehmer die Durchschaubarkeit beider Bedienoberflächen besser als die weiblichen Versuchsteilnehmer. Zudem favorisierten sowohl die weiblichen als auch die männlichen Versuchsteilnehmer die dunkle Bedienoberfläche.

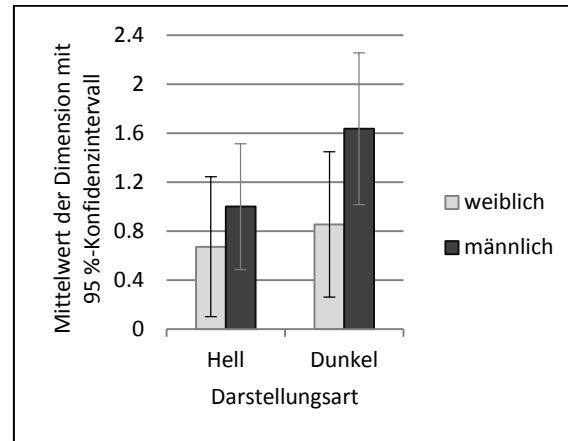


Abbildung 5-45: Mittelwerte und 95 %-Konfidenzintervalle der Dimension *Durchschaubarkeit* bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen

Die Auswertung der Dimension *Effizienz* zeigte keinen nennenswerten Unterschied. Die männlichen Versuchsteilnehmer bewerteten die Effizienz beider Bedienoberflächen besser als die weiblichen Versuchsteilnehmer (Abbildung 5-46). In dieser Dimension favorisierten sowohl die weiblichen als auch die männlichen Versuchsteilnehmer die dunkle Bedienoberfläche.

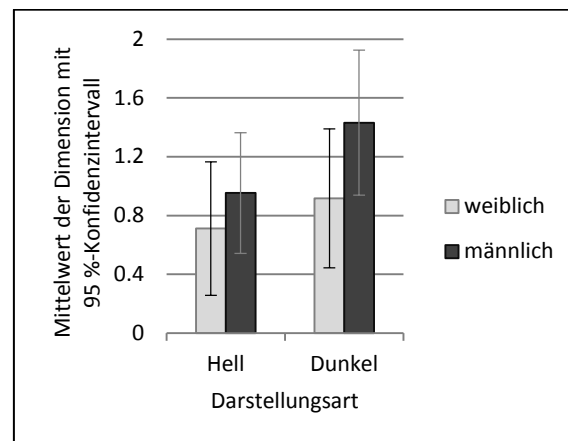


Abbildung 5-46: Mittelwerte und 95 %-Konfidenzintervalle der Dimension *Effizienz* bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen

Abbildung 5-47 zeigt das Ergebnis der Dimension *Steuerbarkeit*. Die männlichen Versuchsteilnehmer bewerteten die Steuerbarkeit beider Bedienoberflächen besser als die weiblichen Versuchsteilnehmer. Zudem wurde in dieser Dimension ebenfalls eine Tendenz der Geschlechter festgestellt. Die weiblichen Versuchsteilnehmer präferierten die helle Bedienoberfläche und die männlichen Versuchsteilnehmer die dunkle Bedienoberfläche.

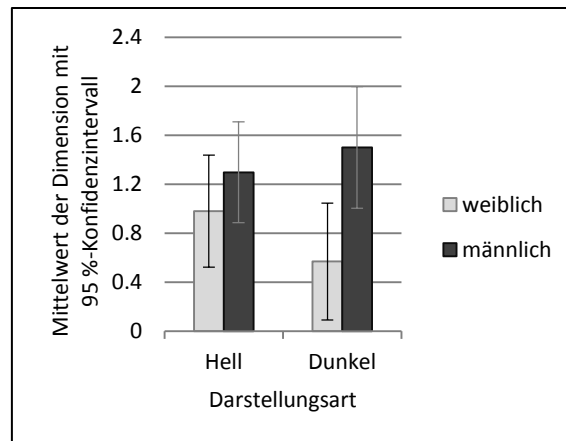


Abbildung 5-47: Mittelwerte und 95 %-Konfidenzintervalle der Dimension *Steuerbarkeit* bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen

Auch in der Dimension *Stimulation* ergab sich ein ähnliches Bild. Die männlichen Versuchsteilnehmer bewerteten die Stimulation beider Bedienoberflächen besser als die weiblichen Versuchsteilnehmer (Abbildung 5-48). Hier wurde ebenfalls eine Tendenz bezogen auf das Geschlecht der Versuchsteilnehmer nachgewiesen. Die weiblichen Versuchsteilnehmer präferierten die helle Bedienoberfläche und die männlichen Versuchsteilnehmer favorisierten die dunkle Bedienoberfläche.

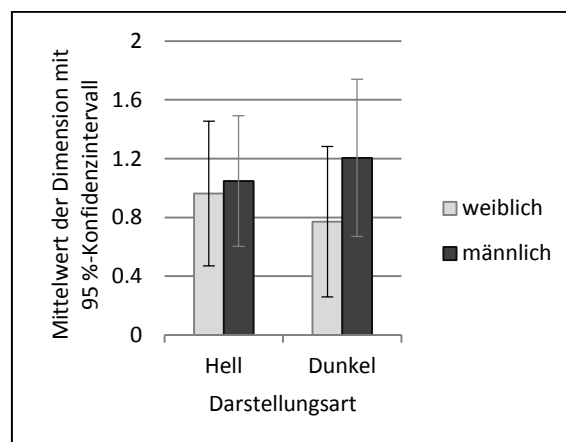


Abbildung 5-48: Mittelwerte und 95 %-Konfidenzintervalle der Dimension *Stimulation* bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen

Am auffälligsten war das Ergebnis der Dimension *Originalität*. Die Originalität der Bedienoberflächen wurde in beiden Versuchsgruppen sehr gering wahrgenommen – wie Abbildung 5-49 zeigt. Eine statistische Erklärung für diesen Umstand wurde nicht gefunden. Die Untersuchung des Zusammenhangs zwischen den unabhängigen Variablen lieferte zumindest eine qualitative Erklärung für die schlechte Bewertung der hellen Bedienoberfläche. Die männlichen Versuchsteilnehmer bewerteten die Dimension *Originalität* der hellen Bedienoberfläche negativ. Dieser Umstand und die, im Vergleich zur Beurteilung der dunklen Bedienoberfläche, geringe Bewertung durch die weiblichen Versuchsteilnehmer, führten zu dem schlechten Ergebnis.

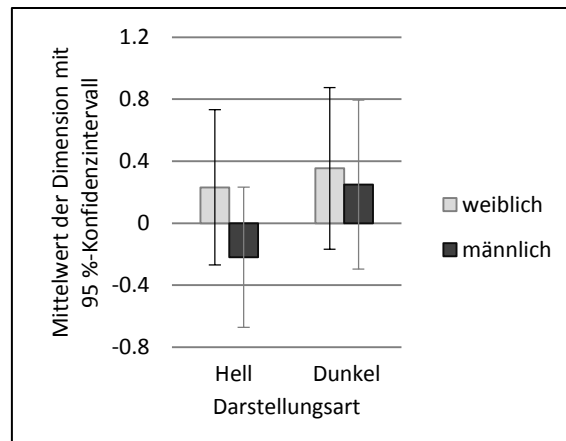


Abbildung 5-49: Mittelwerte und 95 %-Konfidenzintervalle der Dimension *Originalität* bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen

5.2.3.2 Ergebnisse bezogen auf die psychische Beanspruchung

Bei der psychischen Beanspruchung wurden in der Studie signifikante Unterschiede festgestellt. Sowohl die farbliche Gestaltung (Kapitel 5.2.1.2) als auch das Geschlecht der Versuchsteilnehmer (Kapitel 5.2.2.2) wirkten sich auf die Ergebnisse des Items 4 und des Items 7 signifikant aus.

Bei den Antworten des Items 4 wurde ein signifikanter Zusammenhang zwischen den unabhängigen Variablen festgestellt ($F(1, 46) = 4.808$, $p = .033$). Die Effektstärke lag bei $\eta^2 = .095$.⁸⁸ Abbildung 5-50 verdeutlicht den Zusammenhang grafisch. Die weiblichen Versuchsteilnehmer bewerteten die grafische Gestaltung der Bedienoberflächen grundsätzlich schlechter als die männlichen Versuchsteilnehmer. Hinzu kam, dass die Differenz zwischen den Geschlechtern bei der dunklen Bedienoberfläche größer war als bei der hellen Bedienoberfläche.

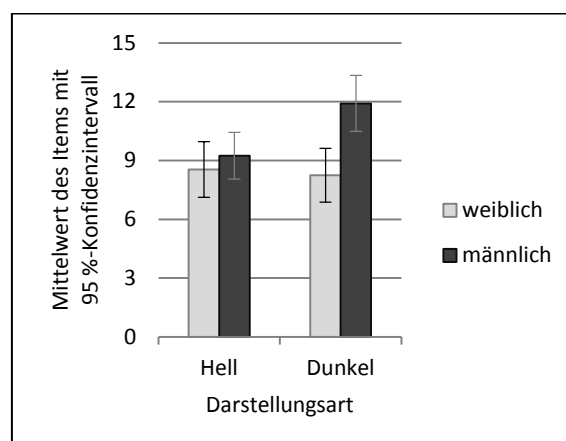


Abbildung 5-50: Mittelwerte und 95 %-Konfidenzintervalle des Items 4 bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen

⁸⁸ Der Effekt ist gering bzw. schwach, wenn $\eta^2 \geq .01$ ist, mittel bzw. moderat bei $\eta^2 \geq .06$ und groß bzw. stark, wenn der Effekt größer als $\eta^2 = .14$ ist [RFH+14a].

Da der Unterschied bezogen auf die farbliche Gestaltung bei Item 4 nur mit dem Mann-Whitney-U-Test nachgewiesen werden konnte und dieser nur marginal signifikant war, wurde der Zusammenhang zwischen den unabhängigen Variablen nicht gewertet und vernachlässigt. Es wurde nur der Einfluss des Geschlechts der Versuchsteilnehmer auf die Ergebnisse des Items 4 akzeptiert. Somit hatte das Geschlecht der Versuchsteilnehmer einen signifikanten Einfluss auf die Bewertung der farblichen Gestaltung der Bedienoberfläche.

Item 7 zeigte keinen signifikanten Zusammenhang. Der Vollständigkeit halber wurde das Item ebenfalls grafisch ausgewertet. Abbildung 5-51 zeigt, dass die weiblichen Versuchsteilnehmer das Erlernen der Bedienung grundsätzlich anstrengender empfanden als die männlichen Versuchsteilnehmer. Hinzu kam, dass die Anstrengung bei der dunklen Bedienoberfläche insgesamt geringer war.

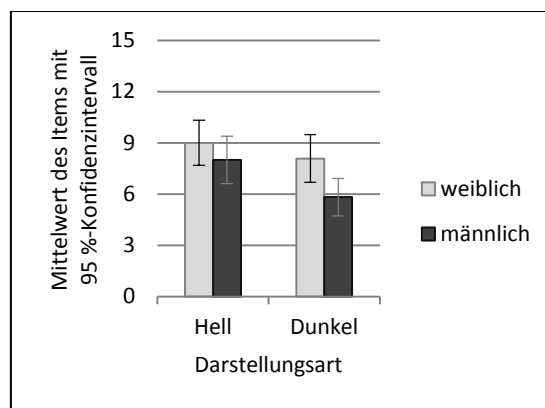


Abbildung 5-51: Mittelwerte und 95 %-Konfidenzintervalle des Items 7 bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen

Eine grafische Auswertung der übrigen Items ergab, bis auf Item 5, keine weiteren Auffälligkeiten. Die Bewertung des Kontrasts der beiden Bedienoberflächen zeigte eine qualitative Tendenz. Sowohl die weiblichen als auch die männlichen Versuchsteilnehmer bewerteten das Kontrastverhältnis der dunklen Bedienoberfläche besser (Abbildung 5-52). Zudem zeigte die Auswertung, dass das Kontrastverhältnis der hellen Bedienoberfläche von den weiblichen Versuchsteilnehmern geringfügig besser bewertet wurde.

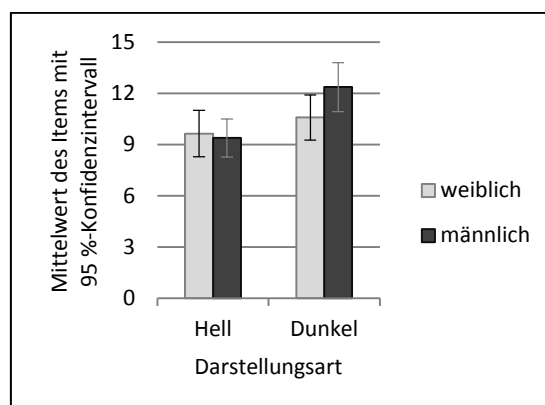


Abbildung 5-52: Mittelwerte und 95 %-Konfidenzintervalle des Items 5 bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen

5.2.3.3 Ergebnisse bezogen auf die Befindlichkeit

Bei der Befindlichkeit wurden in der Dimension *RU* einige signifikante Unterschiede nachgewiesen. Zum einen trat ein Effekt in der Vorher-Nachher-Messung der Gruppe mit der dunklen Bedienoberfläche auf (Kapitel 5.2.1.3.3) und zum anderen wurden zwei Effekte bezogen auf das Geschlecht der Versuchsteilnehmer festgestellt (Kapitel 5.2.2.3.3). Der erste Effekt wurde in der Nachher-Messung zwischen den Geschlechtern festgestellt und der zweite in der Vorher-Nachher-Messung in der Gruppe mit den weiblichen Versuchsteilnehmern.

In den eben beschriebenen Fällen konnte ein Zusammenhang zwischen den unabhängigen Variablen nachgewiesen werden. In der Vorher- und in der Nachher-Messung der Dimension *RU* bestand eine Abhängigkeit zwischen den Variablen. Die Wechselwirkung konnte in der Vorher-Messung als signifikant ($F(1, 48) = 5.955, p = .018$) und in der Nachher-Messung als marginal signifikant ($F(1, 48) = 3.003, p = .090$) nachgewiesen werden. Die Effektstärke in der Vorher-Messung war mittel bis stark ($\eta^2 = .110$) und in der Nachher-Messung klein bis mittel ($\eta^2 = .059$). Der Zusammenhang in der Vorher-Messung wurde nicht weiter betrachtet, da kein Einfluss der Studie auf die Eingangswerte der Versuchsteilnehmer bestand.⁸⁹ Zusammenfassend wurde demnach eine Wechselwirkung bei den Ergebnissen der Nachher-Messung gefunden.

Im Fall der Befindlichkeit waren die Ergebnisse jedoch nicht eindeutig. Es konnte nicht nachvollzogen werden, woher die signifikanten Unterschiede kamen. Die weiblichen Versuchsteilnehmer zeigten nach der Eingewöhnungsphase eine größere Unruhe und wirkten gestresster⁹⁰. Das wurde als mögliche Erklärung für die signifikanten Effekte gewertet. Augenscheinlich wurden die weiblichen Versuchsteilnehmer durch die Eingewöhnung stärker gefordert – was die Ergebnisse in Kapitel 5.2.2.3.3 zeigten. Aus diesem Grund wurde davon ausgegangen, dass der signifikante Unterschied in der Dimension *RU* durch das Geschlecht der Versuchsteilnehmer verursacht wurde.

5.2.3.4 Ergebnisse bezogen auf die Bedienungssicherheit

Bei der Bedienungssicherheit wurden ausschließlich Einflüsse des Geschlechts der Versuchsteilnehmer festgestellt. Demnach konnten auch keine Abhängigkeiten zwischen den unabhängigen Variablen nachgewiesen werden (Anhang G4).

Eine weitere grafische Auswertung ergab eine qualitative Auffälligkeit bei der Differenzbetrachtung zwischen dem ersten und dem letzten Messpunkt (MP4/MP1). Abbildung 5-53 zeigt das Ergebnis. Die weiblichen Versuchsteilnehmer zeigten in der Gruppe mit der hellen

⁸⁹ Des Weiteren wurde in der Vorher-Messung der Dimension *GS* ein Zusammenhang zwischen den unabhängigen Variablen nachgewiesen. Wie auch in der Dimension *RU*, bestand kein Einfluss auf die Eingangswerte der Versuchsteilnehmer. Deshalb wurde dieser Zusammenhang nicht weiter betrachtet. Eine grafische Auswertung ergab keine Auffälligkeiten.

⁹⁰ Wie in Kapitel 5.1.3.4.2 beschrieben, eignet sich der MDBF auch zu einer indirekten Beanspruchungsmessung.

Bedienoberfläche einen größeren Gesamtzuwachs an Bedienungssicherheit als in der Gruppe mit der dunklen Bedienoberfläche. Bei den männlichen Versuchsteilnehmern war es umgekehrt. Hier zeigte sich eine Tendenz bezogen auf das Geschlecht. Die männlichen Versuchsteilnehmer zeigten einen größeren Gesamtsicherheitszuwachs bei der Verwendung der dunklen Bedienoberfläche und die weiblichen Versuchsteilnehmer zeigten diesen Zuwachs bei der Verwendung der hellen Bedienoberfläche.

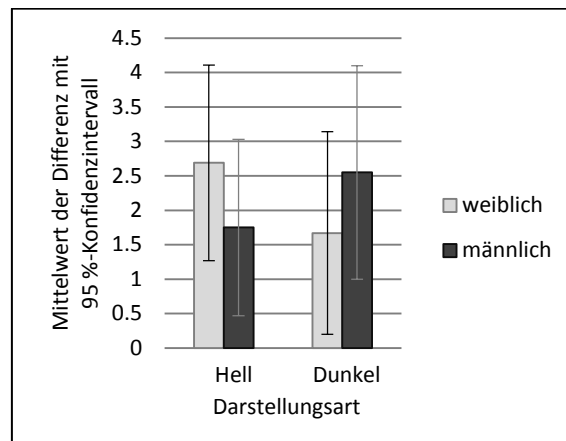


Abbildung 5-53: Mittelwerte und 95 %-Konfidenzintervalle des Gesamtsicherheitszuwachses (MP4/MP1) bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen

Die qualitative Auswertung der Bedienungssicherheit im Messpunkt vor dem Hauptversuch (MP4), dass die weiblichen Versuchsteilnehmer in beiden Gruppen ihre Bedienungssicherheit schlechter einschätzten als die männlichen. Auffällig war in diesem Zusammenhang, dass sich die weiblichen Versuchsteilnehmer in der Gruppe mit der hellen Bedienoberfläche sicherer fühlten als in der Gruppe mit der dunklen Bedienoberfläche (Abbildung 5-54).

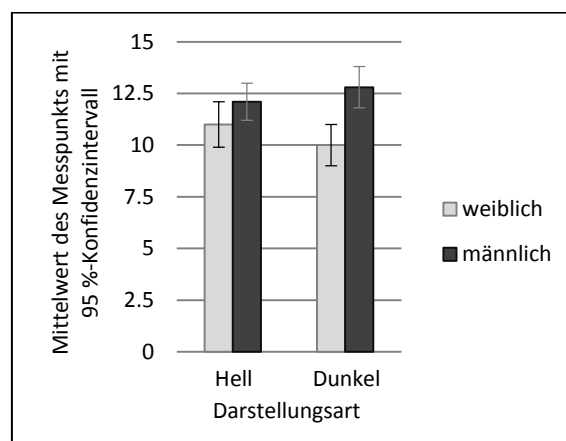


Abbildung 5-54: Mittelwerte und 95 %-Konfidenzintervalle des Messpunkts 4 bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen

5.2.3.5 Ergebnisse bezogen auf die vordefinierten Leistungskennzahlen

Bei den vordefinierten Leistungskennzahlen wurden ausschließlich Einflüsse der farblichen Gestaltung der Bedienoberfläche festgestellt. Demnach konnten auch keine Abhängigkeiten zwischen den unabhängigen Variablen nachgewiesen werden (Anhang G4).

Eine weitere grafische Auswertung ergab keine Auffälligkeiten. Sowohl die weiblichen als auch die männlichen Versuchsteilnehmer erzielten in der Gruppe mit der dunklen Bedienoberfläche bei allen sechs Leistungskennzahlen bessere Ergebnisse. Die Versuchsteilnehmer in der Gruppe mit der hellen Bedienoberfläche schnitten hier insgesamt schlechter ab. Die Untersuchung der Einzelereignisse bei den Fehlerentdeckungszeiten ergab Auffälligkeiten, die weder durch die farbliche Gestaltung (Kapitel 5.2.1.5.2) noch durch das Geschlecht der Versuchsteilnehmer (Kapitel 5.2.2.5.2) erklärt werden konnten. Aus diesem Grund wurden diese Unregelmäßigkeiten zusätzlich grafisch ausgewertet. Die Ergebnisse werden in den folgenden Abbildungen dargestellt (Abbildung 5-55 bis Abbildung 5-60).

Die Fehlerentdeckungszeiten der Unregelmäßigkeit *11_SS_S23* waren in der Gruppe mit der dunklen Bedienoberfläche und bei den weiblichen Versuchsteilnehmern größer. Da die Anzahl der weiblichen Versuchsteilnehmer in der Gruppe mit der dunklen Bedienoberfläche größer war, wurde hier ein Einfluss der Versuchsgruppenszusammensetzung bzw. des Geschlechts vermutet. Abbildung 5-55 zeigt das Ergebnis der Untersuchung bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen.

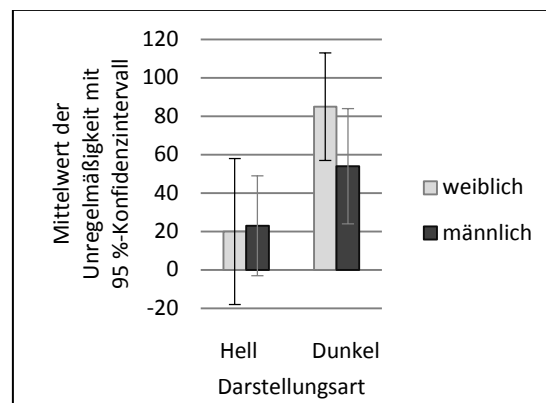


Abbildung 5-55: Mittelwerte und 95 %-Konfidenzintervalle der Unregelmäßigkeit *11_SS_S23* bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen

Die weiblichen Versuchsteilnehmer in der Gruppe mit der dunklen Bedienoberfläche benötigten für das Erkennen der Unregelmäßigkeit mehr Zeit als die weiblichen Versuchsteilnehmer in der Gruppe mit der hellen Bedienoberfläche. Da auch die männlichen Versuchsteilnehmer in der Gruppe mit der dunklen Bedienoberfläche um einiges länger brauchten als die männlichen Versuchsteilnehmer in der Gruppe mit der hellen Bedienoberfläche, wurde ein Einfluss der Versuchsgruppenszusammensetzung ausgeschlossen. Dieser Umstand, die zeitliche Einordnung der Unregelmäßigkeit – gerade am Anfang des Versuchs – und das unterschiedliche Verhältnis der Fehlerentdeckungszeit in den Gruppen, bestätigten die Vermutung aus Kapitel 5.2.1.5.2. Hierbei handelte es sich um eine systematische Abweichung.

Die Untersuchung der Unregelmäßigkeit *17_GFM_W02* zeigte einen Einfluss der farblichen Gestaltung unabhängig vom Geschlecht (Abbildung 5-56). Die Analyse ergab, dass die männlichen Versuchsteilnehmer länger für das Erkennen der Unregelmäßigkeit benötigten und die Fehlerentdeckungszeiten in der Gruppe mit der hellen Bedienoberfläche im Mittel größer war als in der Gruppe mit der dunklen Bedienoberfläche. Das Ergebnis wurde jedoch dadurch relativiert, dass, wie in Kapitel 5.2.1.5.2 und Kapitel 5.2.2.5.2 beschrieben, die Unregelmäßigkeit in der Gruppe mit der hellen Bedienoberfläche fünfmal gar nicht erkannt wurde – viermal davon von den männlichen Versuchsteilnehmern. Aus diesem Grund wurde zwar ein Unterschied zwischen den beiden Bedienoberflächen festgestellt, dieser zeigte jedoch nur eine Tendenz bezogen auf die farbliche Gestaltung.

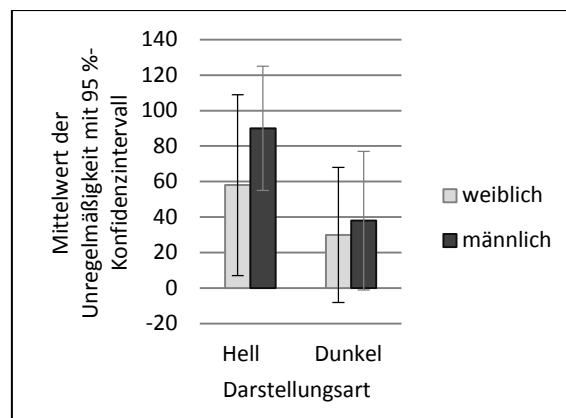


Abbildung 5-56: Mittelwerte und 95 %-Konfidenzintervalle der Unregelmäßigkeit *17_GFM_W02* bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen

Bei der Untersuchung der Unregelmäßigkeit *18_ZLB_S23* wurde in den vorhergehenden Kapiteln festgestellt, dass die Fehlerentdeckungszeiten in der Gruppe mit der dunklen Bedienoberfläche und bei den männlichen Versuchsteilnehmern größer waren. Die Analyse der Wechselwirkung zwischen beiden unabhängigen Variablen bestätigte diese Ergebnisse – wie die Abbildung 5-57 zeigt. Es konnte ein Einfluss der farblichen Gestaltung und des Geschlechts der Versuchsteilnehmer nachgewiesen werden. Sowohl die weiblichen als auch die männlichen Versuchsteilnehmer favorisierten die dunkle Bedienoberfläche.

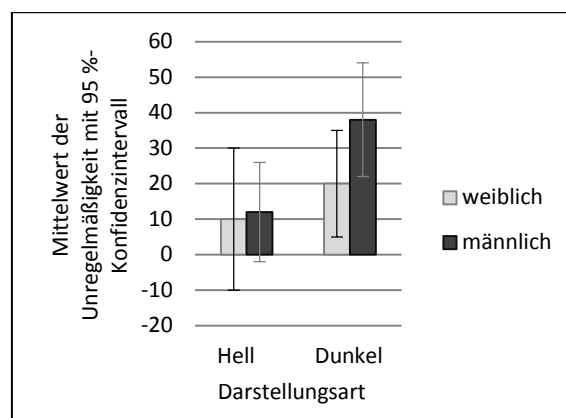


Abbildung 5-57: Mittelwerte und 95 %-Konfidenzintervalle der Unregelmäßigkeit *18_ZLB_S23* bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen

Die Fehlerentdeckungszeiten der Unregelmäßigkeit *19_ZLB_S22* waren in der Gruppe mit der hellen Bedienoberfläche und bei den männlichen Versuchsteilnehmern größer. Abbildung 5-58 zeigt das Ergebnis der Untersuchung bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen.

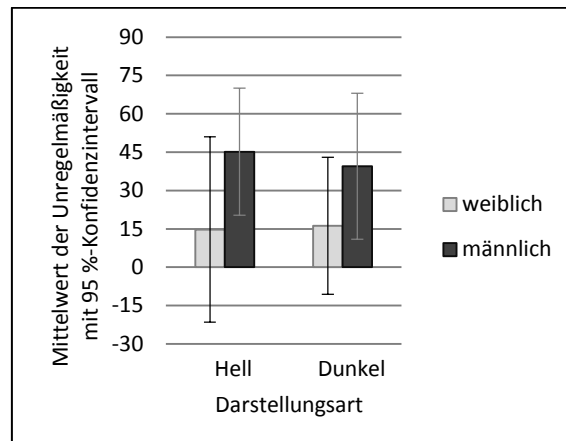


Abbildung 5-58: Mittelwerte und 95 %-Konfidenzintervalle der Unregelmäßigkeit *19_ZLB_S22* bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen

Da die Anzahl der männlichen Versuchsteilnehmer in der Gruppe mit der hellen Bedienoberfläche größer war, zeigte sich hier ein Einfluss der Versuchsgruppenszusammensetzung bzw. des Geschlechts der Versuchsteilnehmer. Die weiblichen Versuchsteilnehmer benötigten in beiden Gruppen ähnlich viel Zeit für das Erkennen der Unregelmäßigkeit, wohingegen die männlichen Versuchsteilnehmer in beiden Versuchsgruppen mehr Zeit für das Erkennen der Unregelmäßigkeit aufbrachten.

Die Ergebnisse der Fehlerentdeckungszeiten der Unregelmäßigkeit *20_ZLB_S23* zeigten, dass das Erkennen dieser Unregelmäßigkeit in der Gruppe mit der hellen Bedienoberfläche und bei den weiblichen Versuchsteilnehmern länger dauerte. Die Untersuchung der Wechselwirkung bestätigte dies jedoch nur zum Teil – wie Abbildung 5-59 zeigt.

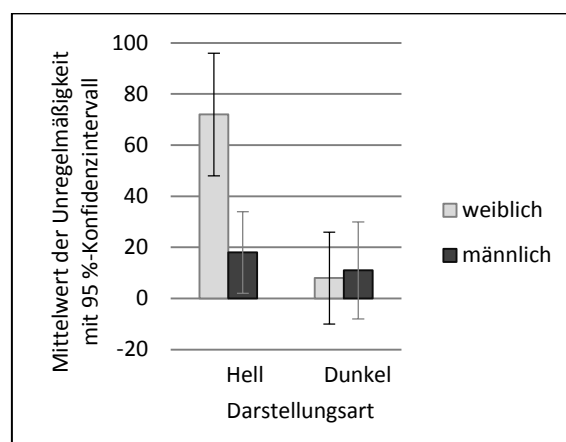


Abbildung 5-59: Mittelwerte und 95 %-Konfidenzintervalle der Unregelmäßigkeit *20_ZLB_S23* bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen

Ausschließlich in der Gruppe mit der hellen Bedienoberfläche benötigten die weiblichen Versuchsteilnehmer sehr viel länger, um die Unregelmäßigkeit zu erkennen. Hierbei handelte es sich vermutlich um einen ungewollten Einfluss durch das nicht ausgewogene Verhältnis zwischen männlichen und weiblichen Versuchsteilnehmern. In der Gruppe mit der hellen Bedienoberfläche war die Anzahl der weiblichen Versuchsteilnehmer sehr klein ($N < 10$).⁹¹ Aus diesem Grund beeinflusste die Gruppenaufteilung das Ergebnis vermutlich negativ und erzeugte eine Auffälligkeit, welche im Nachgang nicht erklärt werden konnte. Die Auffälligkeit war nicht aussagekräftig und wurde deshalb nicht gewertet. Demnach konnte für die dargestellte Unregelmäßigkeit 20_ZLB_S23 weder ein Einfluss der farblichen Gestaltung noch ein Einfluss des Geschlechts der Versuchsteilnehmer nachgewiesen werden.

Abschließend wurde die letzte auffällige Unregelmäßigkeit 25_ZLB_S01 bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen untersucht. Abbildung 5-60 zeigt das Ergebnis der Untersuchung. Das Erkennen der Unregelmäßigkeit 25_ZLB_S01 dauerte in der Gruppe mit der hellen Bedienoberfläche länger als in der Gruppe mit der dunklen Bedienoberfläche. Zudem zeigte sich ein Einfluss des Geschlechts der Versuchsteilnehmer. Die weiblichen Versuchsteilnehmer benötigten in beiden Gruppen mehr Zeit für das Erkennen dieser Unregelmäßigkeit.

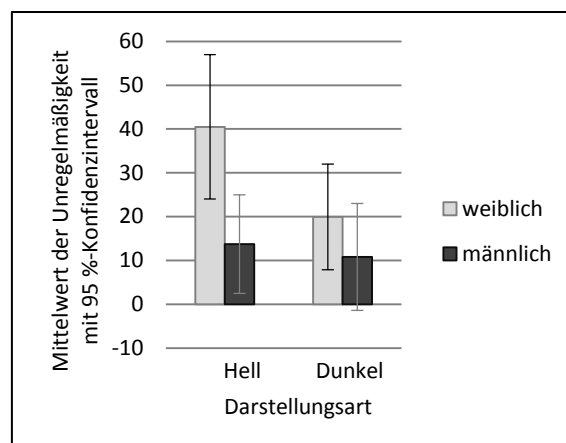


Abbildung 5-60: Mittelwerte und 95 %-Konfidenzintervalle der Unregelmäßigkeit 25_ZLB_S01 bezogen auf die Wechselwirkung zwischen den unabhängigen Variablen

⁹¹ siehe hierzu Kapitel 5.2.1.5, Abbildung 5-26

5.3 Diskussion

Im Folgenden werden die Ergebnisse im Detail diskutiert. Dabei dienen die in Kapitel 5.1.2.1 aufgestellten Hypothesen als Gliederung für die Diskussion des Einflusses der farblichen Gestaltung. Tabelle 5-14 fasst die Hypothesen noch einmal übersichtlich zusammen und gibt einen Überblick über die Hypothesenbewertung nach der Auswertung der Ergebnisse.

Tabelle 5-14: Übersicht der aufgestellten Hypothesen

Nr.	Hypothese	Bewertung
1	Die farbliche Gestaltung der Bedienoberfläche beeinflusst die wahrgenommene User Experience, die empfundene psychische Beanspruchung, die Befindlichkeit, die subjektive Wahrnehmung der Bedienungssicherheit und die erfassten Leistungskennzahlen.	✓
2	Die User Experience, die psychische Beanspruchung und die Befindlichkeit der Versuchsteilnehmer haben einen Einfluss auf die Beurteilung der Bedienungssicherheit. Im Einzelnen wird davon ausgegangen, dass sich hohe Werte bei der wahrgenommenen User Experience, eine geringe Beanspruchung während des Erlernens und eine gute Befindlichkeit nach der Eingewöhnungsphase positiv auf die Bedienungssicherheit auswirken.	✓
3	Eine hohe Bewertung der User Experience und der Bedienungssicherheit sowie eine gute Befindlichkeit, schlagen sich in einer besseren Bedienereffizienz nieder.	✓
4	Versuchsteilnehmer, die sich sicher fühlen, d. h. ihre Bedienungssicherheit selbst hoch einschätzen, verwenden weniger Zeit mit der Überwachung des Bedienbereichs und erzielen ein besseres Ergebnis in der Nebenaufgabe, wobei sich die einzelnen Fehlerentdeckungszeiten durch die geringere Überwachung nicht signifikant erhöhen.	✓
5	Die dunkle Bedienoberfläche erhält, im Vergleich zur hellen Bedienoberfläche, eine bessere Bewertung bei der wahrgenommenen User Experience. Sie wird als attraktiver, gebrauchstauglicher und stimulierender empfunden.	✓
6	Die Versuchsteilnehmer in der Gruppe mit der dunklen Bedienoberfläche sind aktiver und benötigen dadurch weniger Zeit, um eine Unregelmäßigkeit zu erkennen.	✓
Legende: ✓ = Hypothese verifiziert ✓ = Hypothese qualitativ bestätigt ✗ = Hypothese falsifiziert		

Da nicht alle Ergebnisse signifikant waren, konnten die Hypothesen nicht direkt verifiziert bzw. falsifiziert werden. Aus diesem Grund wurde der dritte Status *Hypothese qualitativ bestätigt* eingeführt. Dieser Status zeigt, dass die Hypothese teilweise akzeptiert bzw. teilweise abgelehnt werden konnte. Im weiteren Verlauf werden die Ergebnisse im Detail diskutiert.

5.3.1 Einfluss der farblichen Gestaltung

5.3.1.1 Ergebnis Hypothese 1

Mit *Hypothese 1* wurde erwartet, dass die farbliche Gestaltung der Bedienoberfläche einen Einfluss auf die untersuchten Variablen hat. Diese Annahme lässt sich theoretisch dadurch begründen, dass in [SHL04] gezeigt wurde, dass verschiedenfarbig gestaltete Bedienoberflächen derselben Anwendung, durchaus zu unterschiedlichen Beurteilungen durch den Benut-

zer führen können. Die Ergebnisse der Studie stehen hier im Einklang mit dieser Feststellung. Es zeigte sich ein Einfluss der farblichen Gestaltung der Bedienoberfläche auf die abhängigen Variablen *User Experience*, *psychische Beanspruchung*, *Befindlichkeit*, *Bedienungssicherheit* und *Bedienereffizienz*⁹².

Der Vergleich der Bedienoberflächen ergab, dass die dunkle Bedienoberfläche bei der Bewertung der UX bessere Ergebnisse erzielte als die helle Bedienoberfläche. Dieses Verhältnis setzte sich bei fast allen weiteren abhängigen Variablen fort. So war die empfundene psychische Beanspruchung während des Erlernens der Simulation in der Gruppe mit der dunklen Bedienoberfläche ebenfalls geringer. Hinzu kam, dass die farbliche Gestaltung der dunklen Bedienoberfläche insgesamt als angenehmer wahrgenommen wurde. Des Weiteren wurde das Erkennen von Unregelmäßigkeiten hier als leichter empfunden. Die dunkle Bedienoberfläche beeinflusste die Stimmungslage der Versuchsteilnehmer zudem kaum (Dimension *GS*) und führte zu einer positiven Veränderung der Befindlichkeit in der Dimension *WM*. Diese Einschätzungen bzw. Bewertungen wirkten sich positiv auf die Leistungsfähigkeit und die Handlungssicherheit der Versuchsteilnehmer in der Gruppe mit der dunklen Bedienoberfläche aus. Auch hier zeigten sich insgesamt bessere Ergebnisse. Ausschließlich der Gesamtzuwachs an empfundener Bedienungssicherheit war in der Gruppe mit der hellen Bedienoberfläche höher, wobei hier ein Einfluss des Geschlechts bei der Bewertung nachgewiesen werden konnte. Genauso verhielt es sich mit der Befindlichkeit in der Dimension *RU*. Die Teilnehmer in der Gruppe mit der dunklen Bedienoberfläche waren nach der Eingewöhnungsphase unruhiger als die Versuchsteilnehmer in der Gruppe mit der hellen Bedienoberfläche. Dies konnte ebenfalls auf das Geschlecht der Versuchsteilnehmer zurückgeführt werden.

Zusammenfassend bestätigten die Ergebnisse zu einem Großteil die Vorteilhaftigkeit der dunklen Bedienoberfläche. Dies korrespondierte mit den Ergebnissen aus Kapitel 4.2.4. Hier wurde gezeigt, dass die dunkle Bedienoberfläche barrierefreier ist als die helle Bedienoberfläche. Gerade die erreichten Konformitätsstufen der Farbkonstellationen waren bei dieser Bewertung von Bedeutung, da für eine hohe Konformitätsstufe ein hohes Kontrastverhältnis unabdingbar ist. Wie im Folgenden noch gezeigt wird, ist das hohe Kontrastverhältnis für weitere wichtige Effekte verantwortlich. Demnach waren die Vorgaben des Konzepts für eine barrierefreie Informationstechnik dazu geeignet, eine erste Einschätzung über die Benutzerfreundlichkeit der Stellwerksbedienoberfläche zu treffen.

5.3.1.2 Ergebnis Hypothese 2

Um die Wirkrichtungen der eben beschriebenen Einflüsse zu bestimmen, wurde im nächsten Schritt untersucht, in welchem Zusammenhang die einzelnen abhängigen Variablen zueinander standen. Hierfür wurde zu Beginn die Wirkung der User Experience (UX), der psychischen Beanspruchung und der Befindlichkeit der Versuchsteilnehmer auf die Bedienungssicherheit

⁹² Welche in der Arbeit durch die beiden Aspekte *Leistungsfähigkeit* und *Handlungssicherheit* operationalisiert wurde.

untersucht. Mit *Hypothese 2* wurde erwartet, dass sich hohe Werte bei der wahrgenommenen UX, eine geringe Beanspruchung und eine gute Befindlichkeit positiv auf die Bedienungssicherheit auswirken. Als Bezugspunkt für die Beurteilung der Bedienungssicherheit wurde der letzte Messpunkt (MP4) vor dem Hauptversuch herangezogen. Dieser Messpunkt stand in einem zeitlichen Bezug zu den restlichen erfassten Variablen, wodurch die Vergleichbarkeit der Variablen gewahrt blieb.

Die Ergebnisse der Bedienoberflächenstudie bestätigten diese Annahme. Es konnte ein positiver Zusammenhang zwischen der Bewertung der UX und der Bedienungssicherheit nachgewiesen werden. So führte eine gute Bewertung der UX zu einer besseren Beurteilung der Bedienungssicherheit. Vor allem die Dimensionen *Durchschaubarkeit* ($r_s(50) = .286$, $p = .044$), *Effizienz* ($r_s(50) = .368$, $p = .009$) und *Steuerbarkeit* ($r_s(50) = .410$, $p = .003$) zeigten hier signifikante Effekte. Die Effektstärke war moderat. Ebenso korrelierte die Dimension *Stimulation* positiv mit der Bedienungssicherheit. Dieser Effekt war nicht signifikant und nur schwach, wurde aber dennoch akzeptiert, da der festgestellte Trend mit früheren Studien übereinstimmte – siehe hierzu [Wan93]. Die Attraktivität und die Dimension *Originalität* hatten keinen Einfluss auf die Beurteilung der eigenen Bedienungssicherheit.

Des Weiteren zeigten sich interessante Effekte zwischen der Bewertung der UX und der psychischen Beanspruchung. Es konnte ein starker signifikanter Zusammenhang zwischen der Dimension *Durchschaubarkeit* und der psychischen Beanspruchung (Item 1–3, Item 7), in positiver Wirkrichtung, nachgewiesen werden. Hinzu kam eine moderate signifikante Abhängigkeit zwischen der Bewertung der Dimension *Steuerbarkeit* und der psychischen Beanspruchung (Item 1–3, Item 7) – ebenfalls in positiver Richtung.

In diesem Zusammenhang wurde auch die Korrelation zwischen der UX und der Befindlichkeit untersucht. Es zeigte sich, dass bis auf die Dimension *Originalität* alle Dimensionen des UEQ signifikant mit den Dimensionen des MDBF in positiver Richtung korrelierten. Vor allem die gefundenen Effekte zwischen den Dimensionen *Durchschaubarkeit*, *Effizienz* und *Steuerbarkeit* und der Dimension *GS* sowie zwischen der Dimension *Stimulation* und der Aktivierungsdimension *WM* waren hierbei stark. In der Dimension *Originalität* konnten keine Abhängigkeiten festgestellt werden. Dies war vermutlich darauf zurückzuführen, dass diese Dimension von den Versuchsteilnehmern insgesamt sehr schlecht bewertet wurde.

Zwischen der psychischen Beanspruchung und der Bedienungssicherheit konnte auch ein Zusammenhang in erwarteter Wirkrichtung nachgewiesen werden. Die Bewertung des Items 1 und des Items 7 des Fragebogens zur Beurteilung des Trainings korrelierte negativ mit der Bedienungssicherheit. Die Effektstärke war hierbei moderat [Item 1 ($r_s(50) = -.388$, $p = .005$); Item 7 ($r_s(50) = -.371$, $p = .008$)]. Das bedeutet, je geringer die Beanspruchung während des Erlernens bewertet wurde, desto besser fiel die Beurteilung der eigenen Bedienungssicherheit vor dem Hauptversuch aus. Die getrennte Untersuchung, bezogen auf die geistigen Aspekte (Item 2) und die körperlichen Aspekte (Item 3), zeigte dasselbe Bild. Auch hier konnten

signifikante Korrelationen nachgewiesen werden [Item 2 ($r_s(50) = -.314, p = .026$); Item 3 ($r_s(50) = -.311, p = .028$)].

Die Untersuchung der Zusammenhänge zwischen der psychischen Beanspruchung und der Befindlichkeit zeigte qualitative Korrelationen. Die psychische Beanspruchung korrelierte negativ mit allen Dimensionen des MDBF. Dieser Zusammenhang konnte weiterverfolgt werden, da auch alle Dimensionen des MDBF mit der Bedienungsicherheit positiv korrelierten. Vor allem in den Dimensionen *GS* und *RU* wurden signifikante Effekte nachgewiesen. [*GS* ($r_s(50) = .502, p < .001$), *RU* ($r_s(50) = .608, p < .001$)]. Die Korrelation der Bedienungsicherheit mit der Dimension *RU* wurde jedoch nicht akzeptiert, da im Rahmen der Auswertung der Daten festgestellt wurde, dass bei der Bewertung der Dimension *RU* das Geschlecht der Versuchsteilnehmer einen Einfluss hatte.

Zusammenfassend führte vor allem eine hohe pragmatische Qualität der Bedienoberfläche zu einer besseren Beurteilung der Bedienungsicherheit. Die Attraktivität und die hedonische Qualität der Bedienoberfläche wirkten sich gar nicht bzw. nur qualitativ auf die Beurteilung der Bedienungsicherheit aus. Des Weiteren konnte gezeigt werden, dass sich die gute Durchschaubarkeit und die gute Steuerbarkeit der dunklen Bedienoberfläche positiv auf die empfundene Beanspruchung auswirkten und die wahrgenommene UX der Bedienoberfläche insgesamt einen nicht unerheblichen Einfluss auf die Befindlichkeit der Versuchsteilnehmer hatte. Weiter stützten die Ergebnisse die Aussage, dass eine niedrige Beanspruchung die Einschätzung der Bedienungsicherheit in positiver Richtung förderte und zu einer besseren Befindlichkeit der Versuchsteilnehmer führte. Die bessere Befindlichkeit beeinflusste die Beurteilung der Bedienungsicherheit ebenfalls positiv.

5.3.1.3 Ergebnis Hypothese 3

Es wurde bereits gezeigt, dass sich hohe Werte bei der wahrgenommenen UX, eine geringe Beanspruchung während des Erlernens und eine gute Befindlichkeit positiv auf die Bedienungsicherheit auswirkten. Mit *Hypothese 3* wurde erwartet, dass sich hohe Werte bei der wahrgenommenen UX, eine geringe Beanspruchung, eine gute Befindlichkeit und eine hohe Bedienungsicherheit auch in einer besseren Leistungsfähigkeit und in einer besseren Handlungssicherheit niederschlagen.

Bezogen auf die UX konnte die Hypothese nur teilweise bestätigt werden. Es konnte ein schwacher qualitativer Effekt zwischen der Bewertung der pragmatischen Qualität und der Gesamtfehlerentdeckungszeit nachgewiesen werden. Eine gute pragmatische Qualität führte zu einer geringeren Gesamtfehlerentdeckungszeit ($r_s(40) = -.124, p = .447$). Der Ursprung dieses Zusammenhangs konnte auf die Durchschaubarkeit der Bedienoberfläche zurückgeführt werden. So korrelierte die Bewertung der Durchschaubarkeit signifikant negativ mit den Ergebnissen der Gesamtfehlerentdeckungszeit ($r(40) = -.348, p = .028$). Hinzu kam, dass eine gute Durchschaubarkeit zu einer geringeren Gesamtüberwachungszeit führte

($r(37) = -.362$, $p = .028$). Diese Erkenntnisse stützten die Aussage nach [MoTh10], wonach sich eine gute Usability positiv auf die Leistungsbereitschaft ausüben kann.

Auch die Untersuchung der psychischen Beanspruchung förderte einen signifikanten Effekt zutage. Die empfundene Anstrengung (Item 7) korrelierte mit der Handlungssicherheit. So führte eine hohe Anstrengung zu weniger Hilfshandlungen ($r_s(40) = -.348$, $p = .028$) und zu einer größeren Anzahl an Bedienfehlern ($r_s(40) = .427$, $p = .006$). Das heißt im Umkehrschluss, dass das Risiko von Bedienfehlern mit steigender Beanspruchung zunimmt. Dieses Ergebnis ist besonders für die Arbeit in der Rückfallebene interessant, da hier die Arbeitsbelastung um einiges höher sein kann als im Regelbetrieb.

Bezogen auf den Einfluss der Befindlichkeit wurde erwartet, dass sich eine gute Befindlichkeit positiv auf die Leistungsfähigkeit und die Handlungssicherheit der Versuchsteilnehmer auswirkt. Hierbei konnten zwei qualitative Effekte nachgewiesen werden. Zum einen hatte die Dimension *GS* einen positiven Einfluss auf die Leistungsfähigkeit und die Handlungssicherheit der Versuchsteilnehmer und zum anderen beeinflusste die Dimension *WM* die Fehlererkennung und die Anzahl der gemachten Bedienfehler in positiver Richtung. Eine gute Stimmungslage führte somit zu einer geringeren Gesamtüberwachungszeit ($r_s(37) = -.317$, $p = .056$) und zu einer geringeren Gesamtfehlerentdeckungszeit ($r_s(40) = -.122$, $p = .452$). Ebenso wurde gezeigt, dass Versuchsteilnehmer, welche sich wach fühlten, weniger Zeit zum Erkennen einer Unregelmäßigkeit benötigten ($r_s(40) = -.137$, $p = .399$). Zudem machten sie weniger Bedienfehler ($r_s(40) = -.150$, $p = .356$). Ein ähnlicher Effekt konnte in der Dimension *RU* nachgewiesen werden. Ruhige und entspannte Versuchsteilnehmer machten ebenfalls weniger Bedienfehler ($r_s(40) = -.131$, $p = .421$). Der Großteil dieser Effekte war ausschließlich qualitativer Natur und zeigte nur Tendenzen.

Des Weiteren sollte mit *Hypothese 3* überprüft werden, ob sich eine gute Bedienungssicherheit positiv auf die Leistungsfähigkeit und auf die Handlungssicherheit der Versuchsteilnehmer ausübt. Die Ergebnisse zeigten hier keine signifikanten Zusammenhänge. Es konnten lediglich qualitative Effekte nachgewiesen werden. So führte eine hohe Bedienungssicherheit zu einer geringeren Gesamtfehlerentdeckungszeit ($r_s(40) = -.244$, $p = .128$), zu einer größeren Anzahl an abgefertigten Zugfahrten ($r_s(40) = .210$, $p = .193$) und zu einem besseren Klick-Fehler-Verhältnis (KFV) ($r_s(39) = .148$, $p = .368$). Gerade die negative Korrelation zwischen der Bedienungssicherheit und der Gesamtfehlerentdeckungszeit war von Interesse, da zusätzlich eine Korrelation zwischen dem Item 6 des Fragebogens zur Beurteilung des Trainings, der Bedienungssicherheit und der Gesamtüberwachungszeit nachgewiesen wurde. Die Bewertung der unterstützenden Wirkung der farblichen Gestaltung bei der Fehlererkennung (Item 6) korrelierte signifikant positiv mit der Einschätzung der Bedienungssicherheit ($r_s(40) = .356$, $p = .026$) und signifikant negativ mit der Gesamtüberwachungszeit ($r_s(37) = -.388$, $p = .018$). Die farbliche Gestaltung unterstützte die Fehlererkennung, wodurch sich die Teilnehmer sicherer fühlten und den Bedienbereich weniger überwachen musste. Dies wurde zusätzlich durch die erfasste Handlungssicherheit der Versuchsteilnehmer qualitativ bestätigt. Versuchsteilnehmer, die sich sicher fühlten, verbrachten weniger

Zeit mit der Überwachung des Bedienbereichs ($r_s(37) = -.156, p = .357$), führten mehr Hilfs-handlungen durch ($r_s(40) = .121, p = .458$) und machten weniger Bedienfehler ($r_s(40) = -.270, p = .092$). Da es sich bei den nachgewiesenen Effekten nur um schwache Tendenzen handelte, wurden diese nur eingeschränkt akzeptiert.

Zusammenfassend führte vor allem eine gute Durchschaubarkeit der Bedienoberfläche zu einer geringeren Fehlerentdeckungszeit und zu einer geringeren Gesamtüberwachungszeit der Versuchsteilnehmer. Des Weiteren zeigten die Ergebnisse, dass sich sowohl eine geringe psychische Beanspruchung als auch eine gute Befindlichkeit positiv auf die Anzahl der gemachten Bedienfehler auswirkten. Dieser Effekt wurde auch durch die Beurteilung der eigenen Bedienungssicherheit bestätigt. Das war jedoch nicht weiter überraschend, da mit *Hypothese 2* bereits gezeigt wurde, dass die übrigen abhängigen Variablen einen Einfluss auf die Bedienungssicherheit hatten.

5.3.1.4 Ergebnis Hypothese 4

Mit *Hypothese 3* wurde bereits gezeigt, dass die Bedienungssicherheit einen Einfluss auf das Überwachungsverhalten der Versuchsteilnehmer hatte. Konsequenterweise müssten Versuchsteilnehmer, welche ihre Bedienungssicherheit selbst hoch einschätzen, weniger Zeit mit der Überwachung des Bedienbereichs verbringen (*Hypothese 4*). Gleichzeitig können Sie der Nebenaufgabe eine höhere Aufmerksamkeit widmen ohne dabei einen Leistungseinbruch in der Hauptaufgabe zu riskieren und länger für das Erkennen einer Unregelmäßigkeit zu benötigen.

Die Ergebnisse der Studie bestätigen *Hypothese 4* qualitativ. Die Versuchsteilnehmer, die sich sicher fühlten, verbrachten weniger Zeit mit der Überwachung des Bedienbereichs ($r(37) = -.250, p = .135$). Ebenso konnte der Zusammenhang zwischen der Bedienungssicherheit und dem KfV bestätigt werden. Versuchsteilnehmer, die sich sicher fühlten, erzielten ein besseres Klick-Fehler-Verhältnis. Dabei führten die geringere Beobachtung der Betriebs-situation und die höhere Aufmerksamkeit für die Nebenaufgabe nicht zu einer größeren Gesamtfehlerentdeckungszeit [GÜ ($r(37) = .166, p = .325$); KfV ($r(39) = -.209, p = .201$)]. Somit hatte die Nebenaufgabe keinen negativen Einfluss auf die Hauptaufgabe.

5.3.1.5 Ergebnis Hypothese 5

Mit *Hypothese 5* wurde geprüft, ob die Tendenz der Voruntersuchung auf eine größere Stichprobe übertragen werden kann. Wie auch in der Voruntersuchung, zeigte sich, dass die Versuchsteilnehmer die dunkle Bedienoberfläche in allen drei UX-Qualitäten besser bewerteten. Dieses Ergebnis wurde zusätzlich durch die Items 4–6 des Fragebogens zur Beurteilung des Trainings bestärkt. Die farbliche Gestaltung (Item 4) und das Kontrastverhältnis (Item 5) wurden in der Gruppe mit der dunklen Bedienoberfläche als angenehmer empfunden. Hinzu

kam, dass die unterstützende Wirkung der farblichen Gestaltung bei der Fehlererkennung (Item 6) in der Gruppe mit der dunklen Bedienoberfläche ebenfalls besser bewertet wurde.

In diesem Zusammenhang wurde festgestellt, dass die Bewertung der farblichen Gestaltung und des Kontrasts signifikant mit der Attraktivität und der pragmatischen Qualität in positiver Richtung korrelierte. Anscheinend wirkte sich ein guter Kontrast positiv auf die Akzeptanz des Bedieners aus und führte zu einer besseren Bewertung der Usability. Das wiederum unterstützte den Bediener bei der Fehlererkennung. Die Korrelation zwischen Item 4 und der UX war zu erwarten, da gerade dieser Aspekt mit dem Item gesamthaft erfasst werden sollte. Hierdurch wurde die Verwendbarkeit dieses Items zusätzlich bestätigt.

Neben der Bestätigung der Tendenz aus der Voruntersuchung wurden weitere Effekte nachgewiesen. So führte eine gute farbliche Gestaltung auch zu einer Verbesserung der Handlungssicherheit. Die Bewertung der farblichen Gestaltung korrelierte signifikant mit der Anzahl der Hilfshandlungen ($r_s(40) = .434, p = .005$), mit der Anzahl der Bedienfehler ($r_s(40) = -.371, p = .018$) und mit der Gesamtüberwachungszeit ($r_s(37) = -.456, p = .005$). Ebenso zeigte sich ein Einfluss des Kontrasts auf die erfassten Leistungskennzahlen. Eine gute Bewertung des Kontrasts führte zu einer größeren Anzahl an abgefertigten Zugfahrten ($r_s(40) = .428, p = .006$), zu einer geringeren Gesamtfehlerentdeckungszeit ($r_s(40) = -.329, p = .038$), zu einer größeren Anzahl an durchgeführten Hilfshandlungen ($r_s(40) = .421, p = .007$) und zu einer geringeren Gesamtüberwachungszeit ($r_s(37) = -.456, p = .005$).

5.3.1.6 Ergebnis Hypothese 6

Hypothese 6 bezog sich auf die Aktivierung der Versuchsteilnehmer durch die farbliche Gestaltung. Die Überlegung dahinter war, dass eine ansprechende Bedienoberfläche die Wachheit der Versuchsteilnehmer unterstützt, wodurch sich deren Gesamtfehlerentdeckungszeit verringert. Da die dunkle Bedienoberfläche von den Versuchsteilnehmern in jeder der drei UX-Qualitäten besser bewertet wurde, müssten die Versuchsteilnehmer in dieser Gruppe wacher sein und weniger Zeit für das Erkennen einer Unregelmäßigkeit benötigen.

Die Ergebnisse der Studie bestätigten diese Vermutung. Die Versuchsteilnehmer in der Gruppe mit der dunklen Bedienoberfläche bewerteten die Dimension *WM* nach der Eingewöhnungsphase besser – sie waren demnach wacher. Dies wurde zusätzlich durch eine positive Korrelation der farblichen Gestaltung (Item 4) und des Kontrasts (Item 5) gestützt. Es konnte also gezeigt werden, dass eine als angenehm empfundene farbliche Gestaltung und ein guter Kontrast eine positive Wirkung auf die Aktivierung der Bediener hatten. Dies führte dazu, dass die Versuchsteilnehmer weniger Zeit für das Erkennen einer Unregelmäßigkeit benötigten ($r_s(40) = -.137, p = .399$). Es bestand ein positiver Zusammenhang zwischen der Aktivierung und der Fehlererkennung. Das Ergebnis der Bedienoberflächenstudie steht somit im Einklang mit früheren Studien – siehe hierzu [PfWi86], [PfWi87].

5.3.2 Einfluss des Geschlechts der Versuchsteilnehmer

Neben den Befunden für den Einfluss der farblichen Gestaltung wurden auch geschlechterspezifische Unterschiede bei der Bewertung und während der Versuchsdurchführung festgestellt.

Die weiblichen Versuchsteilnehmer bewerteten, bis auf die Dimension *Originalität*, alle Dimensionen des UEQ schlechter als die männlichen. Hinzu kam, dass signifikante Effekte in der Bewertung der psychischen Beanspruchung und in der Dimension *RU* des MDBF nachgewiesen wurden. Das bedeutete, dass die weiblichen Versuchsteilnehmer eher gestresst waren und dass sie das Erlernen der Simulation als anstrengender empfanden. Ebenso auffällig waren die Ergebnisse bzgl. der Beurteilung der Bedienungssicherheit. Die weiblichen Versuchsteilnehmer schätzten sich selbst grundsätzlich unsicherer ein. Dies spiegelte sich auch in schlechteren Ergebnissen bei den Leistungskennzahlen wider. Hier stimmte demnach die Selbsteinschätzung mit der Realität überein. Der Ursprung dieser Unsicherheit wurde in der größeren Beanspruchung und der Unruhe der weiblichen Versuchsteilnehmer gesehen.

Da sich die Auswertung in der vorliegenden Arbeit ausschließlich auf geschlechterspezifische Auffälligkeiten bezog, wurden weitere Einflussfaktoren der Person nicht betrachtet. Dies könnte als mögliche Begrenzung der Ergebnisse gesehen werden, da neben dem Geschlecht auch weitere persönliche Einflussfaktoren denkbar waren. So könnten Faktoren wie z. B. der Studiengang und das Nutzungsverhalten von Unterhaltungselektronik, einen nicht unerheblichen Einfluss auf den Umgang mit der Simulation gehabt haben. Dies liegt jedoch, mangels empirischer Belege, nur im Bereich der Mutmaßungen.

Interessanter an dieser Stelle war die Untersuchung der geschlechterspezifischen Präferenzen bzgl. der Bedienoberflächen. Die männlichen Versuchsteilnehmer favorisierten die dunkle Bedienoberfläche in allen Bereichen, was sich auch in den Leistungskennzahlen widerspiegelte. Bei den weiblichen Versuchsteilnehmern war das Ergebnis nicht so eindeutig. Sie bewerteten die Beanspruchung während des Erlernens der Simulation bei der hellen Bedienoberfläche geringer, fühlten sich hier wohler und schätzten ihre Bedienungssicherheit höher ein. Hinzu kam, dass die weiblichen Versuchsteilnehmer die helle Bedienoberfläche als attraktiver, steuerbarer und stimulierender empfanden. Auf der anderen Seite bewerteten sie die Durchschaubarkeit und die Effizienz der dunklen Bedienoberfläche besser und erzielten hier auch bessere Ergebnisse bei den Leistungskennzahlen.

Dies mag auf den ersten Blick im Widerspruch zu den bisher gefundenen Effekten in Bezug auf die farbliche Gestaltung stehen. Hier wurde gezeigt, dass die psychische Beanspruchung, die Befindlichkeit und zum Teil auch die Bedienungssicherheit einen positiven Einfluss auf die Leistungsfähigkeit und die Handlungssicherheit hatten. Eine bessere Bewertung der eben genannten abhängigen Variablen müsste im Umkehrschluss auch zu besseren Ergebnissen bei den Leistungskennzahlen führen. Dies ist jedoch nicht zwangsläufig so. Bei der Untersuchung des Einflusses der farblichen Gestaltung wurden ausschließlich die Korrelationen un-

abhängig von Geschlecht und Gruppe über alle Versuchsteilnehmer aufgezeigt. Aus diesem Grund könnten Gruppeneffekte und/oder einzelne Aspekt einen zusätzlichen Einfluss auf die Ergebnisse gehabt haben. Gerade die gute Durchschaubarkeit, der als angenehm bewertete Kontrast und die positive Einschätzung der unterstützenden Wirkung der farblichen Gestaltung bei der Fehlererkennung, könnten eine mögliche Erklärung für die bessere Leistung der weiblichen Versuchsteilnehmer in der Gruppe mit der dunklen Bedienoberfläche sein. Wie die Ergebnisse zeigten, korrelierten diese drei funktionalen Aspekte signifikant mit den Leistungskennzahlen. Sie beschreiben folglich wichtige Einflussgrößen für eine bessere Leistungsfähigkeit und eine gute Handlungssicherheit.

Zusammenfassend zeigte sich ganz klar eine geschlechterspezifische Präferenz. Die männlichen Versuchsteilnehmer tendierten in allen erfassten Bereichen zu der dunklen Bedienoberfläche, wohingegen die weiblichen Versuchsteilnehmer ihre Einschätzung von verschiedenen Faktoren abhängig machten. Sie arbeiteten lieber mit der hellen Bedienoberfläche, bewerteten die dunkle Bedienoberfläche in Bezug auf funktionale Eigenschaften aber besser. Das spiegelte sich dann auch in einer besseren Leistungsfähigkeit und in einer größeren Handlungssicherheit wider. Dies kann für künftige Studien durchaus von Interesse sein und dieser Ansatz sollte weiter verfolgt werden.

5.3.3 Weiterführende Diskussion

An dieser Stelle werden Erkenntnisse aus der Bedienoberflächenstudie diskutiert, welche dem Versuchsleiter während der Auswertung der Daten auffielen bzw. welche die Auswertung der Daten erschwerten.

In Kapitel 2.2.2.2 wurde gezeigt, dass die erfasste Attraktivität die Akzeptanz bzw. die Ablehnung der Benutzer gegenüber dem Produkt global beschreiben kann. Die Ergebnisse der Studie bestätigten diese Aussage. Sowohl die pragmatische als auch die hedonische Qualität korrelierten signifikant positiv mit der Attraktivität. Die Effektstärke zwischen der pragmatischen Qualität und der Attraktivität war moderat ($r_s(52) = .309$, $p = .026$) und stark zwischen der Attraktivität und der hedonischen Qualität ($r_s(52) = .647$, $p < .001$). Demnach wirkte sich eine positive Bewertung der hedonischen Qualität stärker auf die Attraktivität aus als eine positive Bewertung der pragmatischen Qualität – also der Usability. Dieses Ergebnis ist nicht überraschend, da in [ISH09] gezeigt wurde, dass zwischen der Ästhetik⁹³ und der Attraktivität ein positiver Zusammenhang besteht. Dies konnte durch die Ergebnisse der vorliegenden Studie weiter untermauert werden. Ein weiterer interessanter Aspekt war der gefundene Zusammenhang zwischen der pragmatischen und der hedonischen Qualität. Hier konnte ein schwacher positiver Effekt nachgewiesen werden. Eine als ästhetisch empfundene Bedienoberfläche führte zu einer besseren Bewertung der Usability. Das bestätigte die Ergebnisse aus [HaAu05] und bekräftigt die Kernaussage nach [TKI00]: „*What is beautiful is usable*“.

⁹³ ein Aspekt der hedonischen Qualität

Die Untersuchung der Korrelation der einzelnen Leistungskennzahlen zeigte deutlich, dass die erfassten durchsatzbezogenen Bezugsgrößen in der Studie dazu geeignet waren, eine Aussage über die Bedienereffizienz der Versuchsteilnehmer zu treffen. So ergab sich, dass Versuchsteilnehmer mit einer hohen Handlungssicherheit auch leistungsfähiger waren. Das heißt, sie machten weniger Bedienfehler, führten mehr Hilfshandlungen durch und verbrachten weniger Zeit mit der Überwachung des Bedienbereichs. Zudem fertigten sie mehr Zugfahrten ab und benötigten weniger Zeit für das Erkennen einer Unregelmäßigkeit. Ein weiterer interessanter Zusammenhang konnte zwischen der Anzahl ausgeführter Hilfshandlungen und der Anzahl gemachter Bedienfehler nachgewiesen werden. Es zeigte sich, dass geübte Versuchsteilnehmer weniger Bedienfehler machten, obwohl sie mehr Bedienhandlungen, in Form von Hilfshandlungen, durchführten. Für die Praxis bedeutet das, auch wenn die Steuerbereiche größer werden und dadurch die Anzahl der registrierungspflichtigen Hilfshandlungen zunimmt, muss das nicht zwangsläufig zu mehr Bedienfehlern führen. Das gilt nur unter der Voraussetzung, dass gut ausgebildetes und gut trainiertes Bedienpersonal die Hilfshandlungen ausführt. Dieses Ergebnis sollte Bestandteil weiterführender Studien sein.

Der Schwierigkeitsgrad der Gleisfreimeldestörung wurde im Nachgang als zu hoch eingeschätzt. Die Versuchsteilnehmer benötigten stellenweise sehr lang, um die Gleisfreimeldestörung von einer regulären Zugfahrt zu unterscheiden – dies setzte einen sehr guten Trainingsstand voraus. Zudem war die Abfolge der notwendigen Hilfshandlungen sehr umfangreich, wodurch einige Teilnehmer zeitweise sehr stark abgelenkt wurden. Diese Art der Unregelmäßigkeit ist sehr gut dafür geeignet, um Stress und Beanspruchung bei dem Versuchsteilnehmer zu erzeugen, aber nicht zur Untersuchung von Fehlerentdeckungszeiten.

Mit Blick auf die Fehlerentdeckungszeiten, zeigte sich eine Schwäche des verwendeten Eye-trackers. Zu Beginn der Studie wurde angenommen, dass sich die Fehlerentdeckungszeiten im betrieblichen Kontext im Bereich mehrerer Sekunden bewegen würden. Diese Annahme mag in der Realität korrekt sein, erwies sich im Experiment jedoch als hinderlich. Da die Versuchsteilnehmer auf einen begrenzten Bereich fixiert waren, benötigten sie stellenweise sehr wenig Zeit für das Erkennen einer Unregelmäßigkeit – teilweise sogar unter einer Sekunde. Diese Zeiten konnten mit dem verwendeten Eyetracker nicht erfasst werden. Aus diesem Grund muss in weiterführenden Studien ein anderes Messsystem gewählt werden, um belastbarere Ergebnisse zu erhalten. Das verwendete Gerät eignet sich ausschließlich für die Untersuchung des Blickverhaltens der Versuchsteilnehmer.

Ähnlich verhielt es sich mit den erstellten Heatmaps. Durch die wenigen Blickdaten des Eye-trackers, waren die Heatmaps nicht aussagekräftig und demnach auch nicht dazu geeignet, um Unterschiede im Blickverhalten der einzelnen Versuchsteilnehmer zu erkennen. Die Blicke der Versuchsteilnehmer konzentrierten sich hauptsächlich auf den Bildschirmmittelpunkt. Zudem zeigte sich, dass die Versuchsteilnehmer die Gleislage mit ihren Augen abfuhrten und die Signale sowie das Dialogfenster der KF-Bedienung des Öfteren fixierten. Das war nicht weiter verwunderlich, da hier am häufigsten mit wichtigen Informationen zu rechnen war. Weitere Auffälligkeiten konnten nicht festgestellt werden.

Mit dem Headset konnten keine verwendbaren Daten gesammelt werden. Ursprünglich sollte den Versuchsteilnehmern mit diesem Verfahren die Möglichkeit gegeben werden, ihre Gedanken direkt zu verbalisieren. Diese Möglichkeit wurde von den Teilnehmern jedoch nicht genutzt. Die Kommentierung ihrer Handlungen und das laute Aussprechen ihrer Gedanken schien den Versuchsteilnehmern fremd und stellenweise unangenehm zu sein. Aus diesem Grund schwieg der Großteil der Teilnehmer und es standen keine auswertbaren Daten zur Verfügung.

Der Standardisierungsgrad der Befragung zum Designbedienplatz war zu gering. Aus diesem Grund konnten die Antworten nicht kategorisiert werden. Grundsätzlich zeigte sich jedoch, dass die Versuchsteilnehmer die farbliche Gestaltung des Designbedienplatzes als angenehm empfanden und das Design als modern einschätzten. Zudem fanden die Teilnehmer die graue Darstellung der Grundstellung der Gleisfreimeldung übersichtlich und weniger grell. Um diese Ergebnisse quantitativ bestätigen zu können, sollte der Versuch mit dem Designbedienplatz wiederholt werden.

5.3.4 Mögliche versuchsbedingte Einschränkungen

Abschließend wird auf zwei mögliche Begrenzungen der Ergebnisse der vorliegenden Arbeit eingegangen. Zum einen könnten die unterschiedlichen Umgebungsbedingungen einen störenden Einfluss gehabt haben und zum anderen die Tatsache, dass überwiegend Studierende anstelle von Fahrdienstleitern eingesetzt wurden.

Hintergrund der ersten Überlegung war, dass *LAUGWITZ* in ihrer Arbeit konstatiert, dass, wenn Regeln in einem bestimmten Anwendungskontext identifiziert werden sollen, darauf geachtet werden muss, dass die Umgebungsbedingungen den Bedingungen der eigentlichen Anwendung entsprechen [Lau01]. Dieser Forderung konnte in der Studie nur zum Teil Rechnung getragen werden. Es wurde zwar versucht die Anforderungen an die Versuchsteilnehmer so real wie möglich zu gestalten und die Versuchsbedingungen zu standardisieren⁹⁴, letztendlich entsprach die Versuchsumgebung jedoch nur bedingt einem realen Fahrdienstleiterarbeitsplatz. Dieser Umstand begrenzt die ökologische Validität des Experiments zum Teil und schränkt die Übertragbarkeit der Ergebnisse auf die reale Arbeitsumgebung des Fahrdienstleiters möglicherweise ein.

Vor dem Hintergrund, dass die Hauptaufgabe der vorliegenden Arbeit darin bestand wissenschaftliche Grundlagen zu schaffen, diese empirisch zu belegen und für die weitere Forschung aufzubereiten, ist die Forderung nach einer möglichst hohen ökologischen Validität aus Sicht des Verfassers zweitrangig. Es galt in erster Linie fundierte und brauchbare wissenschaftliche Erkenntnisse aus den Ergebnissen abzuleiten. Hierzu leistet die vorliegende Arbeit definitiv ihren Beitrag.

⁹⁴ Beispielsweise wurde, unabhängig von der Tageszeit, immer für gleiche Beleuchtungsverhältnisse gesorgt. Hierdurch sollte diese mögliche Störvariable konstant gehalten werden.

Eine weitere Begrenzung mag darin gesehen werden, dass Studierende anstelle von ausgebildetem Bedienpersonal verwendet wurden. Diese Einschränkung kann zum einen dadurch entkräftet werden, dass die Verwendung von Studierenden im universitären Forschungsalltag zur gängigen Praxis gehört und zum anderen dadurch, dass eine ähnliche Tendenz auch bei erfahrenen Fahrdienstleitern nachgewiesen wurde – siehe hierzu [VAC05]. Die Autoren haben in einem vereinfachten Versuch zwei Bedienoberflächen unter zwei verschiedenen Beleuchtungsstärken durch erfahrenes Bedienpersonal bewerten lassen. Eine der Bedienoberflächen besaß eine positive Polarität und die andere eine negative Polarität im klassischen Sinne⁹⁵. Das Bedienpersonal bevorzugte die Bedienoberfläche mit der negativen Polarität – sowohl bei heller als auch bei dunkler Umgebungsbeleuchtung.

Als möglicher Kritikpunkt wäre die Vorerfahrung der Teilnehmer zu sehen. Die erfahrenen Fahrdienstleiter waren es gewohnt mit dieser Art von Bedienoberfläche zu arbeiten und standen einem neuen Design vermutlich eher kritisch gegenüber. Der Nachweis, dass dem nicht so war, wurde mit der vorliegenden Arbeit erbracht. Auch die unerfahrenen Versuchsteilnehmer in der Bedienoberflächenstudie bevorzugten die Bedienoberfläche mit der „klassischen“ negativen Polarität. Dies ist eine wichtige Erkenntnis, nicht nur aus wissenschaftlicher, sondern auch aus entwicklungstechnischer Sicht.

5.3.5 Zusammenfassung der Abhängigkeiten

Abbildung 5-61 gibt einen Überblick über die gefundenen Abhängigkeiten bezogen auf die farbliche Gestaltung. In der Abbildung werden jedoch nur Zusammenhänge mit mindestens schwacher Effektstärke dargestellt ($r/r_s \geq .10$). Die signifikanten Korrelationen wurden zusätzlich durch einen Punkt gekennzeichnet. Dabei ist die Abbildung 5-61 in drei Ebenen aufgebaut. Auf der ersten Ebene wird die unabhängige Variable, d. h. die farbliche Gestaltung, dargestellt. Die zweite Ebene zeigt die fünf abhängigen Variablen (mit den jeweiligen Dimensionen) und auf der dritten Ebene werden die Korrelationen zwischen den abhängigen Variablen untereinander abgebildet. Da die Richtungen der Einflüsse einem logischen Muster folgen und die Einflüsse stellenweise nur Tendenzen darstellten, wurde auf die zusätzliche Darstellung der Wirkrichtung verzichtet. Das Folgen eines logischen Musters bedeutet in diesem Kontext, dass sich eine positive bzw. negative Entwicklung auch in der logischen und offensichtlichen Richtung widerspiegelt. Beispielsweise wirkte sich eine ansprechende farbliche Gestaltung positiv auf die Bewertung der UX aus. Hieraus ergab sich eine gute Bewertung der Durchschaubarkeit der Bedienoberfläche. Das wiederum führte zu einer geringeren psychischen Beanspruchung und zu einer besseren Befindlichkeit. Dies gilt natürlich auch in umgekehrter Richtung. So führt eine schlechte Durchschaubarkeit der Bedienoberfläche zu einer höheren psychischen Beanspruchung und zu einer schlechteren Befindlichkeit usw.

⁹⁵ Der Wortlaut *im klassischen Sinn* steht in diesem Zusammenhang für eine Bedienoberfläche mit schwarzem Hintergrund und hellen (bunten) Vordergrundfarben – wie es bei dem Experimentalbedienplatz der Fall ist.

farbliche Gestaltung	User Experience	Attraktivität		Befindlichkeit ●	
		Durchschaubarkeit	psychische Beanspruchung ●		
			Befindlichkeit ●		
			Bedienungssicherheit ●		
			Leistungs-kennzahlen	Gesamtfehlerentdeckungszeit ●	
				Gesamtüberwachungszeit ●	
		Effizienz	Befindlichkeit ●		
			Bedienungssicherheit ●		
		Steuerbarkeit	psychische Beanspruchung ●		
			Befindlichkeit ●		
			Bedienungssicherheit ●		
		Stimulation	Befindlichkeit	GS ●	
				WM ●	
	Bedienungssicherheit				
	Originalität	—			
	psychische Beanspruchung		Befindlichkeit		
			Bedienungssicherheit ●		
	Befindlichkeit	Gute-Schlechte-Stimmung (GS)	Bedienungssicherheit ●		
			Leistungs-kennzahlen	abgefertigte Zugfahrten	
				Gesamtfehlerentdeckungszeit	
				KFV	
				Anzahl Hilfshandlungen	
		Gesamtüberwachungszeit			
		Wachheit-Müdigkeit (WM)	Leistungs-kennzahlen	Gesamtfehlerentdeckungszeit	
				Anzahl Hilfshandlungen	
				Bedienfehler	
		Ruhe-Unruhe (RU)	Leistungs-kennzahlen	Anzahl Hilfshandlungen	
				Bedienfehler	
	Bedienungssicherheit		Leistungs-kennzahlen	abgefertigte Zugfahrten	
				Gesamtfehlerentdeckungszeit	
				KFV	
				Anzahl Hilfshandlungen	
				Anzahl Bedienfehler	
				Gesamtüberwachungszeit	
	Leistungs-kennzahlen	abgefertigte Zugfahrten	Gesamtfehlerentdeckungszeit ●		
			Anzahl Hilfshandlungen ●		
			Anzahl Bedienfehler ●		
Gesamtüberwachungszeit ●					
Gesamtfehlerentdeckungszeit		Anzahl Bedienfehler ●			
Klick-Fehler-Verhältnis (KFV)		—			
Anzahl Hilfshandlungen		Anzahl Bedienfehler ●			
Anzahl Bedienfehler		Gesamtfehlerentdeckungszeit ●			
Gesamtüberwachungszeit	abgefertigte Zugfahrten				

Abbildung 5-61: Übersicht der Abhängigkeiten (ohne Darstellung der Wirkrichtung)

6 ZUSAMMENFASSUNG

Die vorliegende Arbeit befasst sich mit der Untersuchung des Einflusses der farblichen Gestaltung der Bedienoberfläche auf den Stellwerksbediener. Dabei waren die zentralen Anliegen der Arbeit, die Wirkung der farblichen Gestaltung auf die abhängigen Variablen *User Experience*, *psychische Beanspruchung*, *Befindlichkeit*, *Bedienungssicherheit* und *Bediener-effizienz*⁹⁶ zu untersuchen, den Zusammenhang der Variablen untereinander empirisch fundiert zu klären und mögliche Gestaltungsvorgaben für zukünftige Bedienoberflächen abzuleiten. Hierzu wurde eine experimentelle Studie durchgeführt, in der zwei Bedienoberflächen einer deutschen Stadtbahnsimulation als Versuchsumgebung dienten. Diese unterschieden sich nur hinsichtlich der farblichen Gestaltung voneinander. Die Hauptaufgabe der Versuchsteilnehmer bestand darin, den Bedienbereich zu überwachen, mögliche Unregelmäßigkeiten zu erkennen und auf diese entsprechend zu reagieren.

Im Folgenden werden die Ergebnisse der Bedienoberflächenstudie zusammengefasst. Dabei liegt der Fokus auf den in Kapitel 3 definierten Zielsetzungen.

Einfluss der farblichen Gestaltung

Die Ergebnisse der Studie zeigten deutliche Unterschiede in der Bewertung der beiden verwendeten Bedienoberflächen. Vor allem der Experimentalbedienplatz, d. h. der Bedienplatz mit der dunklen Bedienoberfläche, wurde von den Versuchsteilnehmern in fast allen Punkten besser bewertet. Hinzu kam, dass die Versuchsteilnehmer, welche diesen Bedienplatz verwendeten, bessere Ergebnisse bei den erfassten Leistungskennzahlen erzielten. Sie waren leistungsfähiger und gleichzeitig auch handlungssicherer. Demnach war das erste Ergebnis der Untersuchung, dass die farbliche Gestaltung der Bedienoberfläche einen Einfluss auf den Stellwerksbediener hatte.

Die identifizierten Unterschiede ließen sich zum größten Teil auf das hohe Kontrastverhältnis der jeweiligen Farbkombinationen und die damit verbundene Bewertung der Attraktivität sowie der pragmatischen Qualität zurückführen. Die Kontrastverhältnisse des Experimentalbedienplatzes waren um einiges höher als die des Originalbedienplatzes. Dies konnte zu Beginn des Experiments durch das Konzept der barrierefreien Informationstechnik gezeigt werden – was als zweites Ergebnis der Untersuchung gesehen wird.

⁹⁶ Welche in der Arbeit durch die beiden Aspekte *Leistungsfähigkeit* und *Handlungssicherheit* operationalisiert wurde.

Das Konzept der barrierefreien Informationstechnik war dazu geeignet, um eine erste Aussage über die Benutzerfreundlichkeit der Bedienoberfläche zu treffen und erste grafische Mängel aufzuzeigen. Vor allem die erreichte Konformitätsstufe war hierbei eine wichtige und aussagekräftige Messgröße, da eine hohe Konformitätsstufe im Zusammenhang mit einer guten Bewertung der farblichen Gestaltung der Bedienoberfläche stand.

Ein weiteres Ergebnis war, dass der User Experience Questionnaire (UEQ) dazu verwendet werden konnte, um die User Experience einer Stellwerksbedienoberfläche zu erfassen. Diese Erkenntnis ist für weiterführende Untersuchungen hilfreich, da die Bewertung der User Experience im Zusammenhang mit zahlreichen positiven Effekten stand. So wirkte sich eine gute User Experience insgesamt positiv auf die Befindlichkeit der Versuchsteilnehmer aus. Zudem senkte eine hohe pragmatische Qualität (Usability) der Bedienoberfläche die empfundene psychische Beanspruchung während des Erlernens der Simulation und erhöhte die subjektive Beurteilung der eigenen Bedienungssicherheit. Das wiederum hatte einen positiven Einfluss auf das Überwachungsverhalten. Die Versuchsteilnehmer verwendeten weniger Zeit mit der Überwachung des Bedienbereichs und benötigten ebenfalls weniger Zeit für das Erkennen einer Unregelmäßigkeit. Hierbei spielte vor allem die Durchschaubarkeit der Bedienoberfläche eine sehr entscheidende Rolle.

Die gefundenen Zusammenhänge sind eine mögliche Erklärung dafür, dass die Versuchsteilnehmer, welche mit dem Experimentalbedienplatz arbeiteten, bessere Ergebnisse erzielten als die Versuchsteilnehmer, welche mit dem Originalbedienplatz arbeiteten. Letztere waren angestrenzter und zeigten eine deutliche Verschlechterung in ihrer Stimmungslage sowie in der Aktivierungsdimension. Dies führte anschließend dazu, dass die Gesamtüberwachungszeit, die Gesamtfehlerentdeckungszeit und die Bedienfehler in dieser Gruppe zunahmen. Aus diesem Grund wäre es denkbar, dass eine längere Benutzung des Originalbedienplatzes zu weiteren Leistungseinbußen und Bedienfehlern führen könnte.

Als konkrete Entwicklungsempfehlung für zukünftige Bedienoberflächen kann in erster Linie abgeleitet werden, dass das dargestellte Kontrastverhältnis eine sehr wichtige Rolle spielt. Es sollte möglichst groß sein. Hierdurch werden vor allem die Attraktivität und die pragmatische Qualität der Bedienoberfläche erhöht. Das senkt wiederum die empfundene Beanspruchung und verbessert die Befindlichkeit. Welche weiteren Faktoren der Darstellungsart einen Einfluss auf die eben genannten Aspekte haben, bleibt an dieser Stelle jedoch offen und muss Bestandteil weiterer Studien sein.

Zusammenfassend wird festgehalten, dass die Erkenntnisse der vorliegenden Arbeit aus wissenschaftlicher Sicht von Bedeutung sind. Sie erscheinen auf den ersten Blick trivial, doch zeigen sie deutlich, dass mit einfachen Mitteln für bessere Arbeitsbedingungen am Fahrdienstleiterarbeitsplatz gesorgt werden kann. Vor allem die Feststellung, dass eine als ästhetisch empfundene Bedienoberfläche zu einer besseren Bewertung der Attraktivität und der Usability führen kann, ist dabei von Interesse.

Einfluss des Geschlechts der Versuchsteilnehmer

Die durchgeführte Untersuchung ergab einen weiteren Faktor, welcher bei der Entwicklung zukünftiger Bedienoberflächen betrachtet werden sollte. Die Ergebnisse der Studie belegen, dass die persönlichen Bedürfnisse bzw. Vorlieben der Versuchsteilnehmer ebenfalls einen Einfluss auf die Bewertung der Bedienoberfläche hatten.

Die männlichen Versuchsteilnehmer favorisierten den Experimentalbedienplatz in allen Bereichen, wohingegen die weiblichen Versuchsteilnehmer ihre Einschätzung von dem jeweiligen Anwendungskontext abhängig machten. Sie arbeiteten lieber mit dem Originalbedienplatz, was sich in einer geringeren psychischen Beanspruchung, einer besseren Befindlichkeit und einer höheren Beurteilung der eigenen Bedienungssicherheit zeigte. Auf der anderen Seite bewerteten die weiblichen Versuchsteilnehmer die Durchschaubarkeit, die Effizienz und den Kontrast des Experimentalbedienplatzes besser, weshalb sie bei der Verwendung dieses Bedienplatzes auch bessere Ergebnisse bei den erfassten Leistungskennzahlen zeigten.

Wodurch dieser geschlechterspezifische Einfluss im Detail zustande kam, konnte nicht geklärt werden. Hier könnten beispielsweise Gruppeneffekte eine Rolle gespielt haben. Unabhängig davon, sollte dieser Effekt in weiterführenden Studien näher spezifiziert und bei der Entwicklung zukünftiger Bedienoberflächen berücksichtigt werden. Funktionale Aspekte sind wichtig, um beispielsweise Unregelmäßigkeiten schneller erkennen zu können, doch gibt es weitere Faktoren durch die die Arbeitsbedingungen und somit auch die Leistungsfähigkeit des Bedieners verbessert werden können. In diesem Kontext ist vor allem von der psychischen Beanspruchung und der Befindlichkeit die Rede.

Persönliche Vorlieben sollten bei der Entwicklung einbezogen werden. Zudem sollten moderne Bedienoberflächen auf die Bedürfnisse der Bediener angepasst werden können, ohne dabei funktionale Aspekte wie die Durchschaubarkeit und den Kontrast zu vernachlässigen. Um dieser Forderung gerecht zu werden, könnte der Fdl in der Zukunft die Möglichkeit bekommen, die Darstellung der Bedienoberfläche aus verschiedenen vorgegebenen Layouts auszuwählen. Dieser personalisierte Bedienplatz könnte ein erster Ansatz sein, um den Fdl wieder aktiver am Prozess teilhaben zu lassen und seine Arbeitsmittel den geänderten Anforderungen entsprechend anzupassen. Das Ziel sollte es sein, dem Fdl die Möglichkeit zu geben, sich auf den wesentlichen Aspekt seiner Arbeit zu konzentrieren – der sicheren Durchführung des Eisenbahnbetriebs.

Operationalisierung der Bedienereffizienz im Stellwerkskontext

Die in der Arbeit erfassten Leistungskennzahlen wurden zusätzlich dazu verwendet, um die Bedienereffizienz im Stellwerk zu operationalisieren. Hierfür wurde ein entsprechendes Modell entwickelt. Als mögliche Einflussgrößen wurden nach [Pac15] das Bedienoberflächendesign und das Situationsbewusstsein des Bedieners gesehen. Beide Einflüsse wurden in der

vorliegenden Arbeit über die Aspekte *Leistungsfähigkeit* und *Handlungssicherheit* beschrieben. Für die Erfassung der Leistungsfähigkeit wurden die Kennzahlen *abgefertigte Zugfahrten*, *Fehlerentdeckungszeit* und *Klick-Fehler-Verhältnis* verwendet und für die Handlungssicherheit die Kennzahlen *Anzahl der Hilfshandlungen*, *Anzahl der Bedienfehler* und *Gesamtüberwachungszeit*.

Der erste Aspekt wurde mit Blick auf den Einfluss des Bedienoberflächendesigns auf die Leistungsfähigkeit betrachtet. Analysiert wurde, ob sich die empfundene Ästhetik positiv auf die Usability auswirkt und ob sich dadurch die Leistungsbereitschaft der Versuchsteilnehmer verbessert. Die Ergebnisse der Studie bestätigten diese Zusammenhänge. Gerade die Durchschaubarkeit der Bedienoberfläche hatte einen positiven Einfluss auf die Leistungsfähigkeit.

Mit dem zweiten Aspekt, der Handlungssicherheit, wurde untersucht, ob das Situationsbewusstsein einen Einfluss auf das Arbeitsverhalten hatte. Die Annahme dabei war, dass die Versuchsteilnehmer, welche die Bedienung der Anlage beherrschten, auch gleichzeitig ein besseres Situationsbewusstsein besaßen. Dieser Zusammenhang konnte durch die Ergebnisse der Studie ebenfalls bestätigt werden. Es zeigte sich, dass Versuchsteilnehmer, welche ihre Bedienungssicherheit selbst hoch einschätzten, auch den Umgang mit der Anlage beherrschten. Das führte dazu, dass sie weniger Zeit mit der Überwachung des Bedienbereichs verbrachten und weniger Bedienfehler machten. Dieses Ergebnis konnte zusätzlich durch das verwendete Mousetracking-Verfahren bestätigt werden. Je sicherer sich die Versuchsteilnehmer fühlten, desto koordinierter waren die Bedienhandlungen. Um der Definition der Handlungssicherheit vollumfänglich gerecht zu werden, wurden die Ergebnisse zusätzlich mit der erfassten Befindlichkeit in Zusammenhang gebracht. Hier zeigte sich, dass die Handlungssicherheit durch eine gute Befindlichkeit unterstützt wird.

Mit dem beschriebenen Ansatz konnte keine allgemeingültige Aussage bzgl. der Bedienereffizienz im Stellwerkskontext formuliert werden. Jedoch bestätigten die Ergebnisse die Einflussfaktoren nach [Pac15] und validierten das Mousetracking-Verfahren als näherungsweise Verfahren, um eine erste Aussage über die Bedienereffizienz treffen zu können. Des Weiteren zeigten die Ergebnisse, dass die verwendeten Leistungskennzahlen dazu geeignet waren, die Aspekte *Leistungsfähigkeit* und *Handlungssicherheit* zu beschreiben und in einer Versuchsumgebung eine Aussage über die Bedienereffizienz der Versuchsteilnehmer treffen zu können.

In erster Linie wurde mit der vorliegenden Arbeit eine empirische Basis geschaffen. Künftigen Studien bleibt es überlassen, Wege zu finden, die gewonnenen Erkenntnisse auf den realen Fahrdienstleiterarbeitsplatz zu übertragen und belastbare Variablen zu definieren, um die Bedienereffizienz im Stellwerk zu erfassen.

LITERATURVERZEICHNIS

- [ÅGK97] Åhsberg, E.; Garnberale, F.; Kjellberg, A.: Perceived quality of fatigue during different occupational tasks Development of a questionnaire. In *International Journal of Industrial Ergonomics* 20 (1997), Nr. 2, S. 121–135
- [ÅhGa98] Åhsberg, E.; Gamberale, F.: Perceived fatigue during physical work – An experimental evaluation of a fatigue inventory. In *International Journal of Industrial Ergonomics* 21 (1998), Nr. 2, S. 117–131
- [And09] Anders, E.: Ein Beitrag zur ganzheitlichen Sicherheitsbetrachtung des Bahnsystems. Dissertation, TU Dresden, Dresden (2009)
- [And99] Andersen, T.: Human reliability and railway safety. In *Proceedings of the 16th ESReDA seminar Safety and Reliability in Transport* (1999), S. 1–14
- [Ape86] Apenburg, E.: Befindlichkeitsbeschreibung als Methode der Beanspruchungsmessung. In *Zeitschrift für Arbeits- und Organisationspsychologie* 30 (1986), S. 3–14
- [ApHä84] Apenburg, E.; Häcker, H.: Untersuchung zur Modifikation der Eigenzustandsskala von Nitsch. Forschungsbericht 3/1984, Bergische Universität Wuppertal, Wuppertal (1984)
- [BaBl14] Barner, D.; Bleicher, I.: Die künftige Bedienphilosophie der DB Netz AG für die Betriebsführung. In *Signal+Dracht* 106 (2014), Nr. 1+2, S. 21–23
- [Bah08] Bahner, J. E.: Übersteigertes Vertrauen in Automation: Der Einfluss von Fehlererfahrungen auf Complacency und Automation Bias. Dissertation, TU Berlin, Berlin (2008)
- [Bai83] Bainbridge, L.: Ironies of Automation. In *Automatica* 19 (1983), Nr. 6, S. 775–779
- [BAuA12] Bundesanstalt für Arbeitsschutz und Arbeitsmedizin: Bildschirmarbeit in Leitwarten – Handlungshilfen zur ergonomischen Gestaltung von Arbeitsplätzen nach der Bildschirmarbeitsverordnung. Dezember 2012
- [BAuA14] Bundesanstalt für Arbeitsschutz und Arbeitsmedizin: Bildschirmarbeit in Leitwarten ergonomisch gestalten. Mai 2014
- [BAV15] Bundesamt für Verkehrs: Bericht über die Sicherheit im öffentlichen Verkehr 2014. <http://www.bav.admin.ch/dokumentation/publikationen/00568/00570/00694/index.html?lang=de>, abgerufen am 29.05.2015
- [Ber81] Bernotat, R.: Anzeigengestaltung. In Schmidtke, H. (Hrsg.): *Lehrbuch der Ergonomie*. Carl Hanser Verlag, München (1981), S. 460–471
- [BHL08] Badke-Schaub, P.; Hofinger, G.; Lauche, K.: Human Factors. In Badke-Schaub, P.; Hofinger, G.; Lauche, K. (Hrsg.): *Human Factors. Psychologie sicheren Handelns in Risikobereichen*. Springer Medizin, Heidelberg (2008), S. 10–25
- [BiSc10] Birbaumer, N.; Schmidt, R. F.: Biologische Psychologie. Springer Medizin, Heidelberg (2010)
- [BKS11] Broschmann, D.; Kuchenbecker, J.; Stilling, J.: Tafeln zur Prüfung des Farbensinnes. Thieme, Stuttgart (2011)
- [BoDö15] Bortz, J.; Döring, N.: Forschungsmethoden und Evaluation. *Für Human- und Sozialwissenschaftler*. Springer, Berlin (2015)
- [BoSc10] Bortz, J.; Schuster, C.: Statistik für Human- und Sozialwissenschaftler. *Mit 163 Tabellen*. Springer Medizin, Berlin (2010)
- [BR16] Bayrischer Rundfunk: "Das ist ein eindeutiger Verstoß – das darf er nicht". <http://www.br.de/nachrichten/mittelfranken/inhalt/hintergrund-fahrerlose-u-bahn-nuernberg-100.html>, abgerufen am 07.01.2017
- [Bra05] Braband, J.: Risikoanalysen in der Eisenbahn-Automatisierung. Eurailpress, Hamburg (2005)

- [Bro65] Brown, I. D.: Effect of a car radio on driving in traffic. In *Ergonomics* 8 (1965), Nr. 4, S. 475–479
- [Bub77] Bubb, H.: *Ergonomie des Mensch-Maschine-Systems*. Habilitationsschrift, TU München, München (1977)
- [Bub84] Bubb, H.: Task analysis in system ergonomics. In Schmidtke, H. (Hrsg.): *Ergonomic Data for Equipment Design*. Plenum Press, New York (1984), S. 31–56
- [Bub92] Bubb, H.: *Menschliche Zuverlässigkeit. Definitionen, Zusammenhänge, Bewertung*. ecomed, Landsberg/Lech (1992)
- [BUND07] Bundesministeriums der Justiz und für Verbraucherschutz: Gesetz zur Gleichstellung von Menschen mit Behinderungen (Behindertengleichstellungsgesetz - BGG). Dezember 2007
- [BUND11] Bundesministeriums der Justiz und für Verbraucherschutz: Verordnung zur Schaffung barrierefreier Informationstechnik nach dem Behindertengleichstellungsgesetz (Barrierefreie Informationstechnik-Verordnung – BITV 2.0). September 2011
- [BWS+12] Balfe, N. et al.: Effects of level of signalling automation on workload and performance. In Wilson, J. R. et al. (Hrsg.): *Rail Human Factors Around the World. Impacts on and of People for Successful Rail Operations*. CRC Press, Boca Raton (2012), S. 404–411
- [CaRe78] Cakir, A.; Reuter, H.-J.: Untersuchungen zur Anpassung von Bildschirmarbeitsplätzen an die physische und psychische Funktionsweise des Menschen. Forschungsbericht, TU Berlin, Berlin (1978)
- [CCM+93] Cabon, P. et al.: Human vigilance in railway and long-haul flight operation. In *Ergonomics* 36 (1993), Nr. 9, S. 1019–1033
- [CCOHS14] Canadian Centre for Occupational Health and Safety: Computer Monitors and Display Colours. <https://www.ccohs.ca/oshanswers/ergonomics/office/vdtcolor.html>, abgerufen am 22.06.2016
- [Chr75] Christ, R. E.: Review and analysis of color coding research for visual displays. In *Human Factors: The Journal of the Human Factors and Ergonomics Society* 17 (1975), Nr. 6, S. 542–570
- [Coh88] Cohen, J.: *Statistical power analysis for the behavioral sciences*. Erlbaum, Hillsdale (1988)
- [Col90] Colin, I.: Gestaltungsmaßnahmen zur Erhöhung von Sicherheit und Zuverlässigkeit. In Hoyos, C. Graf; Zimolong, B. (Hrsg.): *Enzyklopädie der Psychologie, Themenbereich D: Praxisgebiete. Serie III Wirtschafts-, Organisations- und Arbeitspsychologie, Bd. 2: Ingenieurpsychologie*. Hogrefe, Göttingen (1990), S. 346–362
- [Com87] Computerwoche: Trend zur Positivdarstellung endlich in Sicht. <http://www.computerwoche.de/a/trend-zur-positivdarstellung-endlich-in-sicht,1162315>, abgerufen am 24.02.2016
- [CPW78] Coren, S.; Porac, C.; Ward, L. M.: *Sensation and Perception*. Academic Press, New York (1978)
- [DaOs88] Davidoff, J. B.; Ostergaard, A. L.: The role of colour in categorial judgements. In *The Quarterly Journal of Experimental Psychology* 40 (1988), Nr. 3, S. 533–544
- [DBAG13a] Deutsche Bahn AG: Leistungs- und Finanzierungsvereinbarung – Infrastrukturzustands- und -entwicklungsbericht 2013. http://www.eba.bund.de/SharedDocs/Publikationen/DE/Finanzierung/IZB/IZB_2013.pdf?__blob=publicationFile&v=7, abgerufen am 18.06.2015
- [DBAG13b] Deutsche Bahn AG: Ril 482.9001: Signalanlagen bedienen – Allgemeines (A07 Begriffsbestimmungen). September 2013
- [DBAG15] Deutsche Bahn AG: Leistungs- und Finanzierungsvereinbarung – Infrastrukturzustands- und -entwicklungsbericht 2014. http://www.eba.bund.de/SharedDocs/Publikationen/DE/Finanzierung/IZB/IZB_2014.pdf?__blob=publicationFile&v=2, abgerufen am 25.07.2016
- [DHF+01] Dzida, W. et al.: Gebrauchstauglichkeit von Software. *ErgoNorm: Ein Verfahren zur Konformitätsprüfung von Software auf der Grundlage von DIN EN ISO 9241 Teile 10 und 11*. Wirtschaftsverlag NW, Bremerhaven (2001)
- [Dhi07] Dhillon, B. S.: *Human Reliability and Error in Transportation Systems*. Springer, London (2007)

- [DIN00] DIN EN ISO 9241: Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten – Teil 12: Informationsdarstellung (ISO 9241-12:1998). August 2000
- [DIN03] DIN EN 50129: Bahnanwendungen – Telekommunikationstechnik, Signaltechnik und Datenverarbeitungssysteme – Sicherheitsrelevante elektronische Systeme für Signaltechnik. Dezember 2003
- [DIN08] DIN EN ISO 9241: Ergonomie der Mensch-System-Interaktion – Teil 110: Grundsätze der Dialoggestaltung (ISO 9241-110:2006). September 2008
- [DIN11a] DIN EN ISO 9241: Ergonomie der Mensch-System-Interaktion – Teil 210: Prozess zur Gestaltung gebrauchstauglicher interaktiver Systeme (ISO 9241-210:2010). Januar 2011
- [DIN11b] DIN EN ISO 26800: Ergonomie – Genereller Ansatz, Prinzipien und Konzepte (ISO 26800:2011). November 2011
- [DIN12] DIN EN ISO 9241: Ergonomie der Mensch-System-Interaktion – Teil 303: Anforderungen an elektronische optische Anzeigen (ISO 9241-303:2011). März 2012
- [DIN14] DIN EN ISO 6385: Grundsätze der Ergonomie für die Gestaltung von Arbeitssystemen (ISO/DIS 6385:2014, Entwurf). Oktober 2014
- [DIN16] DIN EN 50126-1: Bahnanwendungen – Spezifikation und Nachweis von Zuverlässigkeit, Verfügbarkeit, Instandhaltbarkeit und Sicherheit (RAMS). März 2016
- [DIN99] DIN EN ISO 9241: Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten – Teil 11: Anforderungen an die Gebrauchstauglichkeit – Leitsätze (ISO 9241-11:1998). Januar 1999
- [Elm01] Elms, D. G.: Rail safety. In *Reliability Engineering and System Safety* 74 (2001), Nr. 3, S. 291–297
- [End88] Endsley, M. R.: Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors and Ergonomics Society 32nd Annual Meeting* (1988), S. 97–101
- [End95] Endsley, M. R.: Toward a theory of situation awareness in dynamic systems. In *Human Factors: The Journal of the Human Factors and Ergonomics Society* 37 (1995), Nr. 1, S. 32–64
- [ErSi80] Ericsson, K. A.; Simon, H.: Verbal Reports as Data. In *Psychological Review* 87 (1980), Nr. 3, S. 215–251
- [Fab09] Faber, G.: Die 4. Jet-Generation der Verkehrsflugzeuge – zunehmende Automatisierung. *Technische Entwicklungstrends, didaktische Konsequenzen für die Ausbildung der Piloten und rechtliche Konsequenzen für den Betrieb dieser hochautomatisierten Verkehrsflugzeuge*. Forschungszentrum für Verkehrspilotenausbildung (FHP), Darmstadt (2009)
- [Fig09] Figl, K.: ISONORM 9241/10 und Isometrics: Usability-Fragebögen im Vergleich. In Wandke, H.; Kain, S.; Struve, D. (Hrsg.): *Mensch & Computer 2009. Grenzenlos frei?* Oldenbourg, München (2009), S. 143–152
- [FJP11] Feldmann, F.; Jellentrup, N.; Pelz, M.: Optimierung der Mensch-Maschine-Interaktion durch eine gebrauchstaugliche Systemgestaltung. In *Signal+Draht* 103 (2011), Nr. 1+2, S. 36–40
- [FoWe94] Forstreuter, H.; Weitner von Pein, A.: Verfahrensgesicherte Meldebildanzeige für den FdL-Arbeitsplatz bei der Deutschen Bahn AG. In *Signal+Draht* 86 (1994), Nr. 10, S. 8–12
- [Fra13] Frankfurter Rundschau: Bahn-Mitarbeiter unter Verdacht. <http://www.fr-online.de/verkehr/beinahe-zugunglueck-mainz-bahn-mitarbeiter-unter-verdacht,23914936,23897932.html>, abgerufen am 25.04.2016
- [Gei90] Geiser, G.: Mensch-Maschine-Kommunikation. Oldenbourg, München (1990)
- [GiTi06] Giesa, H.-G.; Timpe, K.-P.: Verlässlichkeit von Mensch-Maschine-Systemen. In Zimolong, B.; Konradt, U. (Hrsg.): *Enzyklopädie der Psychologie, Themenbereich D: Praxisgebiete. Serie III Wirtschafts-, Organisations- und Arbeitspsychologie, Bd. 2: Ingenieurpsychologie*. Hogrefe, Göttingen (2006), S. 603–632
- [HaAu05] Harbich, S.; Auer, S.: Rater Bias: The Influence of Hedonic Quality on Usability Questionnaires. In Costabile, M. F.; Paternò, F. (Hrsg.): *Human-Computer Interaction - INTERACT 2005*. Springer, Berlin (2005), S. 1129–1133

- [HaKr01] Hargutt, V.; Krüger, H.-P.: Die Gefahr der Unterforderung? Selbstregulation der Aktivierung in der Fahrsimulation. *The Danger of Underload? Effort Management in the Driving Simulation*. In VDI-Gesellschaft (Hrsg.): *VDI-Berichte – Der Fahrer im 21. Jahrhundert, Nr. 1613*. VDI Verlag GmbH, Düsseldorf (2001), S. 33–48
- [Han37] Hans, E.: Der Lokomotivführer. In *Das glitzernde Band* (1937), Nr. 7, S. 11
- [HaTi00] Hauß, Y.; Timpe, K.-P.: Automatisierung und Unterstützung in Mensch-Maschine-Systemen. In Timpe, K.-P.; Jürgensohn, T.; Kolrep, H. (Hrsg.): *Mensch-Maschine-Systemtechnik. Konzepte, Modellierung, Gestaltung, Evaluation*. Symposium Publishing, Düsseldorf (2000), S. 41–62
- [HaTr06] Hassenzahl, M.; Tractinsky, N.: User experience – a research agenda. In *Behaviour & Information Technology* 25 (2006), Nr. 2, S. 91–97
- [HBK03] Hassenzahl, M.; Burmester, M.; Koller, F.: AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In Ziegler, J.; Szwillus, G. (Hrsg.): *Mensch & Computer 2003. Interaktion in Bewegung*. Vieweg+Teubner Verlag, Stuttgart (2003), S. 187–196
- [HDP+12] Hinz, A. et al.: Die Stimmung in der deutschen Bevölkerung: Referenzwerte für den Mehrdimensionalen Befindlichkeitsfragebogen MDBF. In *Psychotherapie, Psychosomatik, Medizinische Psychologie* 62 (2012), S. 52–57
- [Hei15] Heise Online: Intel-Chef verabschiedet sich vom bisherigen "Moore's Law". <http://www.heise.de/newsticker/meldung/Intel-Chef-verabschiedet-sich-vom-bisherigen-Moores-Law-2751848.html>, abgerufen am 17.07.2016
- [Her09] Herczeg, M.: Software-Ergonomie. *Theorien, Modelle und Kriterien für gebrauchstaugliche interaktive Computersysteme*. Oldenbourg, München (2009)
- [Heu93] Heuer, H.: Bildschirmarbeit und die Ruhelage des Vergenzsystems. In *Zeitschrift für Experimentelle und Angewandte Psychologie* 40 (1993), Nr. 1, S. 72–102
- [Hey07] Heydt, W.: Entwicklung eines Prototyps zur verfahrensgesicherten Steuerung von Hilfshandlungen bei der Bedienung des elektronischen Stellwerks ALISTER. Diplomarbeit, Universität Stuttgart, Stuttgart (2007)
- [HFES00] Human Factors and Ergonomics Society: Definitions of Human Factors and Ergonomics. <http://www.hfes.org/Web/EducationalResources/HFedefinitionsmain.html>, abgerufen am 24.02.2015
- [HHB02] Hinz, A.; Hessel, A.; Brähler, E.: Der Leipziger Stimmungsbogen. In *Zeitschrift für Differentielle und Diagnostische Psychologie* 23 (2002), Nr. 1, S. 55–65
- [HJL08] Hammerl, M.; Jäber, B.; Lemmer, K.: An integrated model for working environments and rail human factors. In de Waard, D. et al. (Hrsg.): *Human Factors for Assistance and Automation. On the Occasion of the Human Factors and Ergonomics Society, Europe Chapter, Annual Meeting in Braunschweig, October 2007*. Shaker Publishing, Maastricht (2008), S. 415–427
- [HMD+82] Hoyos, C. Graf et al.: Belastung und Beanspruchung bei Steuerungs- und Überwachungstätigkeiten – Untersuchungen in Fahrdienstleitungen der Deutschen Bundesbahn. Forschungsbericht, TU München, München (1982)
- [Hof08] Hofinger, G.: Fehler und Unfälle. In Badke-Schaub, P.; Hofinger, G.; Lauche, K. (Hrsg.): *Human Factors. Psychologie sicheren Handelns in Risikobranchen*. Springer Medizin, Heidelberg (2008), S. 36–55
- [Hol07] Holl, F.: Software-Gestaltung: Farbe auf dem Bildschirm. In *Computer und Arbeit* 15 (2007), Nr. 8–9, S. 9–16
- [HPB+00] Hassenzahl, M. et al.: Hedonic and ergonomic quality aspects determine a software's appeal. In Turner, T.; Szwillus, G. (Hrsg.): *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM Press, New York (2000), S. 201–208
- [HRR74] Hildebrandt, G.; Rohmert, W.; Rutenfranz, J.: 12 and 24 hour Rhythms in error frequency of locomotive drivers and the influence of tiredness. In *International Journal of Chronobiology* 2 (1974), S. 178–180

- [Hub10] Huber, O.: Das psychologische Experiment: Eine Einführung. Hogrefe, Bern (2010)
- [IEA00] International Ergonomics Association: Definitions of Human Factors and Ergonomics. <http://www.iea.cc/whats/index.html>, abgerufen am 05.03.2015
- [IEF13] IEF-Werner GmbH: Benutzerfreundlichkeit ist oberstes Gebot. <http://www.ief-werner.de/news/presseinformationen/benutzerfreundlichkeit-ist-oberstes-gebot/>, abgerufen am 17.07.2016
- [IHS09] Ilmberger, W.; Held, T.; Schrepp, M.: Was macht studiVZ attraktiv? In Wandke, H.; Kain, S.; Struve, D. (Hrsg.): *Mensch & Computer 2009. Grenzenlos frei?* Oldenbourg, München (2009), S. 323–332
- [ISH08] Ilmberger, W.; Schrepp, M.; Held, T.: Cognitive Processes Causing the Relationship between Aesthetics and Usability. In Holzinger, A. (Hrsg.): *HCI and Usability for Education and Work*. Springer, Berlin (2008), S. 43–54
- [ISH09] Ilmberger, W.; Schrepp, M.; Held, T.: Was verursacht den Zusammenhang zwischen Ästhetik und Usability? In Wandke, H.; Kain, S.; Struve, D. (Hrsg.): *Mensch & Computer 2009. Grenzenlos frei?* Oldenbourg, München (2009), S. 383–392
- [ISW94] Igbaria, M.; Schiffman, S. J.; Wieckowski, T. J.: The respective roles of perceived usefulness and perceived fun in the acceptance of microcomputer technology. In *Behaviour & Information Technology* 13 (1994), Nr. 6, S. 349–361
- [JeMe12] Jellentrup, N.; Meyer zu Hörste, Michael: Designing usable interactive systems within the railway domain. *A Human Factors Approach*. In Flammini, F. (Hrsg.): *Railway Safety, Reliability, and Security: Technologies and Systems Engineering*. IGI Global, Hershey (2012), S. 317–326
- [Joh93] Johannsen, G.: Mensch-Maschine-Systeme. Springer, Berlin (1993)
- [Jon12] Jonas, W.: Elektronische Stellwerke bedienen. *Der Regelbetrieb*. Bahn Fachverlag, Berlin (2012)
- [JPS76] Johannsen, G.; Pfendler, C.; Stein, W.: Human performance and workload in simulated landing-approaches with autopilot-failures. In Sheridan, T. B.; Johannsen, G. (Hrsg.): *Monitoring Behavior and Supervisory Control*. Plenum Press, New York (1976), S. 83–96
- [Kar12] Karwowski, W.: The discipline of human factors and ergonomics. In Salvendy, G. (Hrsg.): *Handbook of Human Factors and Ergonomics*. Wiley, New Jersey (2012), S. 3–37
- [KEK+96] Krohne, H. W. et al.: Untersuchung mit einer deutschen Version der "Positive and Negative Affect Schedule" (PANAS). In *Diagnostica* 42 (1996), Nr. 2, S. 139–156
- [Klu06] Kluwe, R. H.: Informationsaufnahme und Informationsverarbeitung. In Zimolong, B.; Konradt, U. (Hrsg.): *Enzyklopädie der Psychologie, Themenbereich D: Praxisgebiete. Serie III Wirtschafts-, Organisations- und Arbeitspsychologie, Bd. 2: Ingenieurpsychologie*. Hogrefe, Göttingen (2006), S. 35–69
- [KMZ+14] Knecht, C. et al.: Vertrauen in die Automatisierung, fehlende Situation Awareness & Fertigkeitsverlust durch automatisierte Systeme – Eine subjektive Einschätzung aus der Linienpilotenperspektive. In Grandt, M.; Schmerwitz, S. (Hrsg.): *Der Mensch zwischen Automatisierung, Kompetenz und Verantwortung (DGLR-Bericht, 2014-01)*. Deutsche Gesellschaft für Luft- und Raumfahrt, Bonn (2014), S. 215–230
- [KoHo14] Kochrow, R.; Hoppe, A.: Gestaltungsempfehlungen für Operatorarbeitsplätze – Ableitungen aus einer Eye-Tracking-Studie. In *13. Symposium Energieinnovation* (2014)
- [Kon13] Konsortium: Neue Generation Signaltechnik. *Sektorweite Initiative zur Sicherung der Zukunftsfähigkeit der Leit- und Sicherungstechnik*. Schlussbericht, o. O. (2013)
- [Kor14] Korte, M.: Jung im Kopf. *Erstaunliche Einsichten der Gehirnforschung in das Älterwerden*. Deutsche Verlags-Anstalt, München (2014)
- [Kra85] Kraiss, K.-F.: Fahrzeug- und Prozeßführung. *Kognitives Verhalten des Menschen und Entscheidungshilfen*. Springer-Verlag (1985)

- [Lau01] Laugwitz, B.: Experimentelle Untersuchung von Regeln der Ästhetik von Farbkombinationen und von Effekten auf den Benutzer bei der Anwendung im Benutzeroberflächendesign. Dissertation, Universität Mannheim, Mannheim (2001)
- [LHS08] Laugwitz, B.; Held, T.; Schrepp, M.: Construction and evaluation of a user experience questionnaire. In Holzinger, A. (Hrsg.): *HCI and Usability for Education and Work*. Springer, Berlin (2008), S. 63–76
- [Lom64] Lomow, B. F.: Ingenieurpsychologie. VEB Deutscher Verlag der Wissenschaften, Berlin (1964)
- [LSH06] Laugwitz, B.; Schrepp, M.; Held, T.: Konstruktion eines Fragebogens zur Messung der User Experience von Softwareprodukten. In Heinecke, A. M.; Paul, H. (Hrsg.): *Mensch & Computer 2006. Mensch und Computer im StrukturWandel*. Oldenbourg, München (2006), S. 125–134
- [LSI+09] Laugwitz, B. et al.: Subjektive Benutzerfreundlichkeit quantitativ erfassen: Erfahrungen mit dem User Experience Questionnaire UEQ. In Brau, H. et al. (Hrsg.): *Usability Professionals 2009*. Fraunhofer Verlag, Stuttgart (2009), S. 220–225
- [Lüd15] Lüdtkke, A.: Wege aus der Ironie in Richtung ernsthafter Automatisierung. In Botthof, A.; Hartmann, E. A. (Hrsg.): *Zukunft der Arbeit in Industrie 4.0*. Springer Vieweg, Berlin (2015), S. 125–146
- [MaBa05] Manzey, D.; Bahner, J. E.: Vertrauen in Automation als Aspekt der Verlässlichkeit von Mensch-Maschine-Systemen. In Karrer, K.; Gauss, B.; Steffens, C. (Hrsg.): *Beiträge zur Mensch-Maschine-Systemtechnik aus Forschung und Praxis. Festschrift für Klaus-Peter Timpe*. Symposion Publishing, Düsseldorf (2005), S. 93–109
- [Mac57] Mackworth, N. H.: Some factors affecting vigilance. In *Advancement of Science* 53 (1957), S. 389–393
- [Mah06] Mahlke, S.: Emotionen als Aspekt des Nutzungserlebens – Implikationen für die Praxis. In Bose-nick, T. et al. (Hrsg.): *Usability Professionals 2006*. Fraunhofer Verlag, Stuttgart (2006), S. 141–144
- [Mah08] Mahlke, S.: User Experience of Interaction with Technical Systems. *Theories, Methods, Empirical Results, and Their Application to the Development of Interactive Systems*. Dissertation, TU Berlin, Berlin (2008)
- [Man08] Manzey, D.: Systemgestaltung und Automatisierung. In Badke-Schaub, P.; Hofinger, G.; Lauche, K. (Hrsg.): *Human Factors. Psychologie sicheren Handelns in Risikobranchen*. Springer Medizin, Heidelberg (2008), S. 307–324
- [MDS+91] Myrtek, M. et al.: Psychophysiologische Untersuchungen zur Belastung und Beanspruchung bei Fahrdienstleitern, Lokomotivführern und Busfahrern. Forschungsbericht, Universität Freiburg, Freiburg i. Breisgau (1991)
- [Mil56] Miller, G. A.: The magical number seven, plus or minus two: some limits on our capacity for processing information. In *Psychological Review* 63 (1956), S. 81–97
- [MLB15] Milius, B.; Lindner, T.; Burkhardt, M.: Usability-Studien im Virtuellen Eisenbahnbetriebslabor des IfEV. In *EI – Der Eisenbahningenieur* 66 (2015), Nr. 6, S. 31–35
- [Moln00] Moray, N.; Inagaki, T.: Attention and complacency. In *Theoretical Issues in Ergonomics Science* 1 (2000), Nr. 4, S. 354–365
- [Moo65] Moore, G. E.: Cramming More Components onto Integrated Circuits. In *Electronics* 38 (1965), Nr. 8, S. 114–117
- [Mos12] Moser, C.: User Experience Design. *Mit erlebniszentrierter Softwareentwicklung zu Produkten, die begeistern*. Springer Vieweg, Wiesbaden (2012)
- [MoTh10] Moshagen, M.; Thielsch, M. T.: Facts of visual aesthetics. In *International Journal of Human-Computer Studies* 68 (2010), Nr. 10, S. 689–709
- [MSF09] Maag, C.; Schmitz, M.; Fröschl, T.: Psychologie des Eisenbahnverkehrs. In Krüger, H.-P. (Hrsg.): *Enzyklopädie der Psychologie, Themenbereich D: Praxisgebiete. Serie VI – Verkehrspsychologie, Bd. 2: Anwendungsfelder der Verkehrspsychologie*. Hogrefe, Göttingen (2009), S. 639–709

- [Mut90] Muthig, K.-P.: Informationsaufnahme und Informationsverarbeitung. In Hoyos, C. Graf; Zimolong, B. (Hrsg.): *Enzyklopädie der Psychologie, Themenbereich D: Praxisgebiete. Serie III Wirtschafts-, Organisations- und Arbeitspsychologie, Bd. 2: Ingenieurpsychologie*. Hogrefe, Göttingen (1990), S. 92–120
- [NaPa04] Naumann, P.; Pachl, J.: Leit- und Sicherungstechnik im Bahnbetrieb. *Fachlexikon*. Tetzlaff Verlag, Hamburg (2004)
- [Neb63] Nebylizyn, W. D.: Über die Zuverlässigkeit der Arbeit des Operators in automatisierten Systemen. In *Probleme und Ergebnisse der Psychologie, Sonderheft Ingenieurpsychologie* (1963), S. 61–71
- [NeTi76] Neumann, J.; Timpe, K.-P.: Psychologische Arbeitsgestaltung. VEB Deutscher Verlag der Wissenschaften, Berlin (1976)
- [Nit71] Nitsch, J. R.: Theorie und Skalierung der Ermüdung. *Eine Studie zum Beanspruchungsproblem*. Dissertation, TU Berlin, Berlin (1971)
- [NiUd76] Nitsch, J.; Udris, I.: Beanspruchung im Sport. *Beiträge zur psychologischen Analyse sportlicher Leistungssituationen*. Limpert Verlag, Bad Homburg (1976)
- [NN05] N. N.: Colour Contrast Check. https://snook.ca/technical/colour_contrast/colour.html#fg=00FF00,bg=B4B4B4, abgerufen am 23.08.2016
- [NN10] N. N.: Helligkeitsrechner für RGB-Farben. <http://www.kurztutorial.info/programme/helligkeit-farbe-berechnen/rgb-bright-rechner.htm>, abgerufen am 15.08.2016
- [NN15] N. N.: User Experience Questionnaire. <http://www.ueq-online.org/>, abgerufen am 31.03.2016
- [NN16a] N. N.: Grundlagen von Korrelation und Kovarianz. <http://support.minitab.com/de-de/minitab/17/topic-library/modeling-statistics/regression-and-correlation/correlation-and-covariance/basics-of-correlation-and-covariance/#sensitivity-of-pearson-s-correlation-coefficient-to-outliers>, abgerufen am 02.05.2016
- [NN16b] N. N.: European Initiative Linking Interlocking Subsystems. <http://eulynx.eu/index.php>, abgerufen am 10.09.2016
- [NN0J] N. N.: Farben umrechnen. <http://www.farbtabelle.at/farben-umrechnen/>, abgerufen am 23.08.2016
- [NoDr86] Norman, D. A.; Draper, S. W.: User centered system design. *New perspectives on human-computer interaction*. Erlbaum, Hillsdale (1986)
- [Nor81] Norman, D. A.: Categorization of action slips. In *Psychological Review* 88 (1981), Nr. 1, S. 1–15
- [OsDa85] Ostergaard, A. L.; Davidoff, J. B.: Some effects of color on naming and recognition of objects. In *Journal of Experimental Psychology* 11 (1985), Nr. 3, S. 579–587
- [Pac13] Pachl, J.: Systemtechnik des Schienenverkehrs. *Bahnbetrieb planen, steuern und sichern*. Springer Vieweg, Wiesbaden (2013)
- [Pac15] Pachl, J.: Ein Ansatz zur Bewertung der Bedienereffizienz in Stellwerken. In *EI – Der Eisenbahningenieur* 66 (2015), Nr. 3, S. 40–43
- [Pac16] Pachl, J.: Zentralisierung sicherheitlich beherrschen. In *EI – Der Eisenbahningenieur* 67 (2016), Nr. 1, S. 3
- [Pac98] Pachl, J.: Anforderungen an die sicherheitsgerechte Visualisierung der Betriebslage. In *Signal+Draht* 90 (1998), Nr. 1+2, S. 5–9
- [Pat92] Patrick, J.: Training: Research and Practice. Academic Press, London (1992)
- [Pet80] Peter, J.-H.: Kann die Sicherheitsfahrschaltung SIFA die Dienstfähigkeit von Triebfahrzeugführern gewährleisten? – Eine psychophysiologische Studie zur Effizienz der SIFA als Vigilanzmonitor und zur Beanspruchung von Fahrzeugführern unter Bedingungen extremer Monotonie. Dissertation, Universität Marburg, Marburg (1980)

- [PWi86] Pfendler, C.; Widdel, H.: Vigilance performance when using colour on electronic displays. In *Perceptual and Motor Skills* 63 (1986), S. 939–944
- [PWi87] Pfendler, C.; Widdel, H.: Zur Beeinflussung der Vigilanzleistung durch Farbe in elektrischen Anzeigen. In *Zeitschrift für Arbeitswissenschaft* 41 (1987), Nr. 1, S. 55–58
- [PIRi84] Plath, H.-E.; Richter, P.: Ermüdung, Monotonie, Sättigung, Streß. *Der BMS-Erfassungsbogen*. Psychodiagnostisches Zentrum, Berlin (1984)
- [PMB14] Piepenbrock, C.; Mayr, S.; Buchner, A.: Positive display polarity is particularly advantageous for small character sizes: implications for display design. In *Human Factors* 56 (2014), Nr. 5, S. 942–951
- [PMM+13] Piepenbrock, C. et al.: Positive display polarity is advantageous for both younger and older adults. In *Ergonomics* 56 (2013), Nr. 7, S. 1116–1124
- [Pot88] Pottgießer, H.: Sicher auf den Schienen. *Fragen zur Sicherheitsstrategie der Eisenbahn von 1825 bis heute*. Birkhäuser, Basel (1988)
- [Pre08] Preuß, E.: Eisenbahnunfälle bei der Deutschen Bahn. *Ursachen, Hintergründe, Konsequenzen*. transpress, Stuttgart (2008)
- [Pro11] Process Vogel: Wie eine höhere Bedienereffizienz des Automatisierungssystems erreicht werden kann. <http://www.process.vogel.de/wie-eine-hoehere-bediener-effizienz-des-automatisierungssystems-erreicht-werden-kann-a-339225/>, abgerufen am 29.03.2016
- [Prü97] Prümper, J.: Der Benutzungsfragebogen ISONORM 9241/10: Ergebnisse zur Reliabilität und Validität. In Liskowsky, R.; Velichkovsky, B. M.; Wüschmann, W. (Hrsg.): *Software-Ergonomie '97*. Springer Vieweg, Wiesbaden (1997), S. 253–262
- [Ras86] Rasmussen, J.: Information Processing and Human-Machine Interaction. *An Approach to Cognitive Engineering*. North-Holland, Amsterdam (1986)
- [Ras87] Rasmussen, J.: The definition of human error and a taxonomy for technical system design. In Rasmussen, J.; Leplat, J.; Duncan, K. (Hrsg.): *New Technology and Human Error (New Technologies and Work)*. Wiley, Chichester (1987), S. 23–30
- [Rau96] Rau, R.: Einzelfallanalysen zur Bewertung von Handlungssicherheit in komplexen, automatisierten Systemen. In *Zeitschrift für Arbeits- und Organisationspsychologie* 40 (1996), Nr. 2, S. 75–86
- [Rea90] Reason, J.: Human Error. Cambridge University Press, Cambridge (1990)
- [Rea94] Reason, J.: Menschliches Versagen. *Psychologische Risikofaktoren und moderne Technologien (aus dem Amerikanischen übersetzt von Joachim Grabowski)*. Spektrum Akad. Verl, Heidelberg (1994)
- [RFH+14a] Rasch, B. et al.: Quantitative Methoden 1. *Einführung in die Statistik für Psychologen und Sozialwissenschaftler*. Springer, Berlin (2014)
- [Ric00] Richter, G.: Psychische Belastung und Beanspruchung. Stress, psychische Ermüdung, Monotonie, psychische Sättigung. Wirtschaftsverlag NW, Bremerhaven (2000)
- [RiHa98] Richter, P.; Hacker, W.: Belastung und Beanspruchung. *Streß, Ermüdung und Burnout im Arbeitsleben*. Roland Asanger Verlag, Heidelberg (1998)
- [RLM02] Rentsch, M.; Liesemeier, B.; Messerschmidt, W.: Entwicklung modularer Führerstände für Triebfahrzeuge des Regionalverkehrs. In *EI – Der Eisenbahningenieur* 53 (2002), Nr. 11, S. 34–37
- [RSP+13] Rauschenberger, M. et al.: Efficient Measurement of the User Experience of Interactive Products. *How to use the User Experience Questionnaire (UEQ)*. In *International Journal of Artificial Intelligence and Interactive Multimedia* 2 (2013), Nr. 1, S. 39–45
- [RTS13] Rauschenberger, M.; Thomaschewski, J.; Schrepp, M.: User Experience mit Fragebögen messen. *Durchführung und Auswertung am Beispiel des UEQ*. In Brau, H. et al. (Hrsg.): *Usability Professionals 2013*. Fraunhofer Verlag, Stuttgart (2013), S. 72–76
- [RuCh08] Rubin, J.; Chisnell, D.: Handbook of Usability Testing. *How to Plan, Design, and Conduct Effective Tests*. Wiley, Indianapolis (2008)

- [S2R16] Shift2Rail: Signalling Innovation Programme 2 – Advanced Traffic Management and Control Systems. *Automatic Train Operation (ATO) (TD 2.2)*. <http://shift2rail.org/research-development/ip2/>, abgerufen am 11.12.2016
- [Sas58] Sasse, H. W.: Streifzug durch die Geschichte der deutschen Signaltechnik. In *Signal+Draht* 43 (1958), Nr. 11+12, S. 208–222
- [Sch75] Schnell, P.: Die Tätigkeit des Fahrdienstleiters im Mensch-Maschine-System "Fahrdienstleiter, Mechanisches Stellwerk". *Ein Beitrag zur quantitativen Analyse mit Hilfe der digitalen Systemsimulation*. Dissertation, Universität Stuttgart, Stuttgart (1975)
- [Sch81a] Schmidtke, H.: Der Leistungsbegriff in der Ergonomie. In Schmidtke, H. (Hrsg.): *Lehrbuch der Ergonomie*. Carl Hanser Verlag, München (1981), S. 105–111
- [Sch81b] Schmidtke, H.: Mentale Beanspruchung durch informatorische Belastung. In Schmidtke, H. (Hrsg.): *Lehrbuch der Ergonomie*. Carl Hanser Verlag, München (1981), S. 131–147
- [Sch99] Schimmack Ulrich: Strukturmodell der Stimmungen: Rückschau, Rundschau und Ausschau. In *Psychologische Rundschau* 50 (1999), S. 90–97
- [ScSc88] Schmidt, L.; Schmidt, G. A.: Belastung und Beanspruchung von Fahrdienstleitern. In *WISO – Wirtschafts- und sozialpolitische Zeitschrift des ISW* 11 (1988), Nr. 4, S. 43–58
- [SDJ+90] Snyder, H. L. et al.: Effect of Image Polarity on VDT Task Performance. In *Proceedings of the Human Factors Society 34th Annual Meeting* (1990), S. 1447–1451
- [She87] Sheridan, T. B.: Supervisory Control. In Salvendy, G. (Hrsg.): *Handbook of Human Factors*. Wiley, New York (1987), S. 1243–1268
- [ShJo76] Sheridan, T. B.; Johannsen, G.: Introduction and summary. In Sheridan, T. B.; Johannsen, G. (Hrsg.): *Monitoring Behavior and Supervisory Control*. Plenum Press, New York (1976), S. 473–482
- [SHL04] Schrepp, M.; Held, T.; Laugwitz, B.: Hedonische Aspekte betriebswirtschaftlicher Software. In Keil-Slawik, R.; Selke, H.; Szwillus, G. (Hrsg.): *Mensch & Computer 2004. Allgegenwärtige Interaktion*. Oldenbourg, München (2004), S. 127–136
- [SSN+94] Steyer, R. et al.: Testtheoretische Analyse des Mehrdimensionalen Befindlichkeitsfragebogens (MDBF). In *Diagnostica* 40 (1994), Nr. 4, S. 320–328
- [SSN+97] Steyer, R. et al.: Der Mehrdimensionale Befindlichkeitsfragebogen (MDBF). Hogrefe, Göttingen (1997)
- [StHo14] St.Pierre, M.; Hofinger, G.: Human Factors und Patientensicherheit in der Akutmedizin. Springer Medizin, Heidelberg (2014)
- [ThMa07] Thüring, M.; Mahlke, S.: Usability, aesthetics and emotions in human–technology interaction. In *International Journal of Psychology* 42 (2007), Nr. 4, S. 253–264
- [THN09] Thielsch, M. T.; Hassenzahl, M.; Nikolaeva, D.: User Experience – aus Sicht der Forschung. In Brau, H. et al. (Hrsg.): *Usability Professionals 2009*. Fraunhofer Verlag, Stuttgart (2009), S. 241–243
- [TKI00] Tractinsky, N.; Katz, A.; Ikar, D.: What is beautiful is usable. In *Interacting with Computers* 13 (2000), Nr. 2, S. 127–145
- [VAC05] Van Laar, D.; Andrew, I.; Cox, M.: Using Visual Layering Methods to Design Rail Network Control Room Displays. In Wilson, J. R. et al. (Hrsg.): *Rail Human Factors. Supporting the Integrated Railway*. Ashgate Publishing, Aldershot (2005), S. 251–261
- [VaDe02] Van Laar, D.; Deshe, O.: Evaluation of a visual layering methodology for colour coding control room displays. In *Applied Ergonomics* 33 (2002), S. 371–377
- [VDI15] VDI: Menschliche Zuverlässigkeit – Ergonomische Forderungen und Methoden der Bewertung, VDI 4006 Blatt 1:2015-03. März 2015
- [VDI89] VDI-Handlungsempfehlung: Sozialverträgliche Gestaltung von Automatisierungsvorhaben. VDI-Hauptgruppe 1989
- [Vol15] Vollrath, M.: Ingenieurpsychologie. *Psychologische Grundlagen und Anwendungsgebiete*. Kohlhammer, Stuttgart (2015)

- [W3C00] The World Wide Web Consortium (W3C): Techniques For Accessibility Evaluation And Repair Tools. <https://www.w3.org/TR/AERT#color-contrast>, abgerufen am 21.09.2016
- [W3C08] The World Wide Web Consortium (W3C): Web Content Accessibility Guidelines 2.0 – Glossary. <https://www.w3.org/TR/WCAG20/#glossary>, abgerufen am 15.08.2016
- [W3C09] The World Wide Web Consortium (W3C): Web Content Accessibility Guidelines 2.0. <https://www.w3.org/TR/WCAG20/#glossary>, abgerufen am 23.08.2016
- [Wan93] Wandmacher, J.: Software-Ergonomie. Walter de Gruyter, Berlin (1993)
- [WCT88] Watson, D.; Clark, L. A.; Tellegen, A.: Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales. In *Journal of Personality and Social Psychology* 54 (1988), Nr. 6, S. 1063–1070
- [WGH+03] Wickens, C. D. et al.: Attentional models of multitask pilot performance using advanced display technology. In *Human Factors: The Journal of the Human Factors and Ergonomics Society* 45 (2003), Nr. 3, S. 360–380
- [WGH96] Willumeit, H.; Gediga, G.; Hamborg, K.-C.: IsoMetrics: Ein Verfahren zur formativen Evaluation von Software nach ISO 9241/10. In *Ergonomie & Informatik* 13 (1996), Nr. 27, S. 5–12
- [W GK+08] Warm, M. et al.: Postoperative Sense of Well-Being and Quality of Life in Breast Cancer Patients Do not Depend on Type of Primary Surgery. In *Onkologie* 31 (2008), Nr. 3, S. 99–104
- [WHB+13] Wickens, C. D. et al.: Engineering psychology and human performance. Pearson, Boston (2013)
- [WHG+01] Wickens, C. D. et al.: Pilot task management: testing an attentional expected value model of visual scanning. Forschungsbericht, University of Illinois, Savoy (2001)
- [Wic84] Wickens, C. D.: Engineering Psychology and Human Performance. Harper Collins Publishers, Illinois (1984)
- [Wic87] Wickens, C. D.: Information processing, decision-making, and cognition. In Salvendy, G. (Hrsg.): *Handbook of Human Factors*. Wiley, New York (1987), S. 72–107
- [Wic99] Wickens, C. D.: Aerospace psychology. In Hancock, P. A. (Hrsg.): *Human Performance and Ergonomics*. Academic Press, San Diego (1999), S. 195–242
- [WiCa12] Wickens, C. D.; Carswell, C. M.: Information processing. In Salvendy, G. (Hrsg.): *Handbook of Human Factors and Ergonomics*. Wiley, New Jersey (2012), S. 117–161
- [Wid90] Widdel, H.: Steuerung und Überwachung industrieller Prozesse. In Hoyos, C. Graf; Zimolong, B. (Hrsg.): *Enzyklopädie der Psychologie, Themenbereich D: Praxisgebiete. Serie III Wirtschafts-, Organisations- und Arbeitspsychologie, Bd. 2: Ingenieurpsychologie*. Hogrefe, Göttingen (1990), S. 396–425
- [Wie07] Wieland, F.: Untersuchungen zum Einsatz vereinfachter Bedienplätze. Diplomarbeit, TU Dresden, Dresden (2007)
- [WiHo00] Wickens, C. D.; Hollands, J. G.: Engineering Psychology and Human Performance. Harper Collins Publishers, New Jersey (2000)
- [WiNo05] Wilson, J. R.; Norris, B. J.: Rail Human Factors: Past, Present and Future. In Wilson, J. R. et al. (Hrsg.): *Rail Human Factors. Supporting the Integrated Railway*. Ashgate Publishing, Aldershot (2005), S. 3–12
- [WiSt83] Wilde, Gerald J. S.; Stinson, J. F.: The monitoring of vigilance in locomotive engineers. In *Accident Analysis & Prevention* 15 (1983), Nr. 2, S. 87–93
- [WLL+04] Wickens, C. D. et al.: An Introduction to Human Factors Engineering. Pearson, Boston (2004)
- [WMC+12] Wilson, J. R.; Mills, A.; Clarke, T. et al.: Rail Human Factors Around the World. *Impacts on and of People for Successful Rail Operations*. CRC Press, Boca Raton (2012)
- [WMD+13] Wilson, J. R.; Mills, A.; Dadashi, N. et al.: Rail Human Factors. *Supporting Reliability, Safety and Cost Reduction*. CRC Press, Boca Raton (2013)

- [WMD+15] SPARK - The Rail Knowledge Hub: Rail Human Factors.
<https://www.sparkrail.org/Pages/SparkWelcome.aspx>, abgerufen am 05.06.2017
- [WNC+05] Wilson, J. R.; Norris, B. J.; Clarke, T. et al.: Rail Human Factors. *Supporting the Integrated Railway*. Ashgate Publishing, Aldershot (2005)
- [WNC+07] Wilson, J. R.; Norris, B. J.; Clarke, T. et al.: People and Rail Systems. *Human Factors at the Heart of the Railway*. Ashgate Publishing, Aldershot (2007)
- [WPM05] Whitlock, A.; Pethick, J.; Mills, A.: Driver vigilance devices. In Wilson, J. R. et al. (Hrsg.): *Rail Human Factors. Supporting the Integrated Railway*. Ashgate Publishing, Aldershot (2005), S. 120–130
- [WyRi08] Wynn, T.; Richardson, J.: Comparison of subjective workload ratings and performance measure of a reference IVIS task. In *Proceedings of the European Conference on Human Centred Design for Intelligent Transport Systems* (2008)
- [ZeSh16] Zenkov, A.; Shipilov, A.: IOGraphica – Turn your routine work into modern art.
<http://iographica.com/>, abgerufen am 06.04.2016
- [Zim06] Zimolong, B.: Gegenstand und Entwicklung der Ingenieurpsychologie. In Zimolong, B.; Konradt, U. (Hrsg.): *Enzyklopädie der Psychologie, Themenbereich D: Praxisgebiete. Serie III Wirtschafts-, Organisations- und Arbeitspsychologie, Bd. 2: Ingenieurpsychologie*. Hogrefe, Göttingen (2006), S. 3–31
- [Zöl15] Zöll, D.: Tablets für Triebfahrzeugführer im Fernverkehr. In *Deine Bahn* 43 (2015), Nr. 7, S. 6–11

ANHANG

A	Einführungspräsentation (exemplarisch Originalbedienplatz)
B	Handout zur Fehlerdiagnose (exemplarisch Originalbedienplatz)
C	eingesetzte Fragebögen
C1	User Experience Questionnaire (UEQ)
C2	Fragebogen zur Beurteilung des Trainings
C3	Mehrdimensionaler Befindlichkeitsfragebogen (Kurzform A)
C4	Mehrdimensionaler Befindlichkeitsfragebogen (Kurzform B)
C5	Fragebogen zur Beurteilung der Bedienungssicherheit
D	Verteilung der Unregelmäßigkeiten
E	Tool zur Erstellung von Heatmaps
F	Bedienoberflächen
G	Ergebnisse der deskriptiven Statistik und der Interferenzstatistik
G1	Ergebnisse User Experience
G2	Ergebnisse psychische Beanspruchung
G3	Ergebnisse Befindlichkeit
G4	Ergebnisse Bedienungssicherheit
G5	Ergebnisse Leistungskennzahlen

Anhang A: Einführungspräsentation (exemplarisch Originalbedienplatz)



Technische
Universität
Braunschweig



Institut für Eisenbahnwesen
und Verkehrssicherung **IfEV**
TU Braunschweig



**Auswirkung der Darstellungsart auf die Stimmung, Belastung
und die Leistungsfähigkeit**

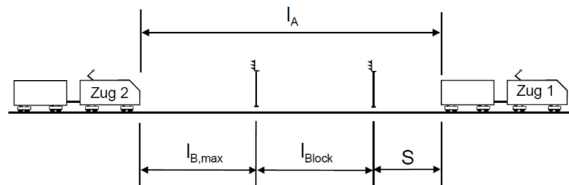
Einführung in die Bedienung des Bedienplatzsystems Vicos® OC 111

Grundlagen des Eisenbahnwesens

Eigenschaften des Spurgeführten Verkehrs:

- Geringe Reibung zwischen Rad und Schiene
 - Lange Bremswege
 - Lokführer kann seinen Bremsweg nicht einsehen
- **Ortsfeste Signale sind nötig um Zugfahrten zu sichern**

Grundlagen des Eisenbahnwesens



Fahren im festen
Raumabstand

- Im spurgeführten Verkehr wird im Regelfall im *festen Bremswegabstand* gefahren

Grundlagen des Eisenbahnwesens

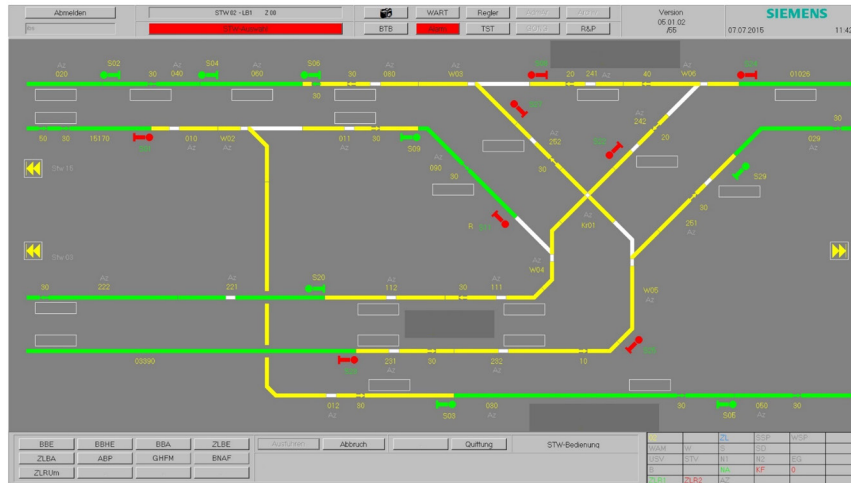


- Signale und Weichen werden vom
Fahrdienstleiter (Fdl), der sich im Stellwerk
befindet, gesteuert



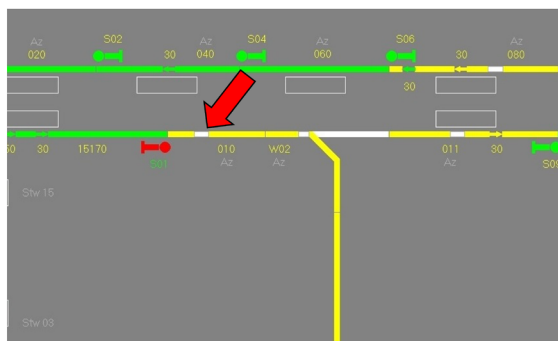
Grundlagen des Eisenbahnwesens

Bedienplatzsystem Vicos OC 111 der Fa. Siemens



Verschlussmelder

- Der Verschchlussmelder trifft Aussage über die Verwendung des Fahrwegelements
- „ausgefüllt“ bedeutet, dass das Element Teil einer Fahrstraße ist und durch diese verschlossen wurde



Eine Fahrstraße ist ein technisch gesicherter Fahrweg für Zug- bzw. Rangierfahrten.

Störungsbetrieb

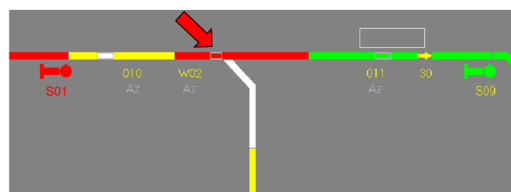
Im Betrieb kann es vorkommen, dass ein Element der Anlage ausfällt. In unserem Beispiel treten drei Arten von Störungen auf:

- Gleisfreimeldestörungen (Gleis zeigt „belegt“ an, obwohl es frei ist),
- Signalstörungen (Ausfall der Grünlampe, Signal kommt nicht mehr auf Fahrt),
- der Zuglenkbetrieb an einzelnen Signalen fällt aus.

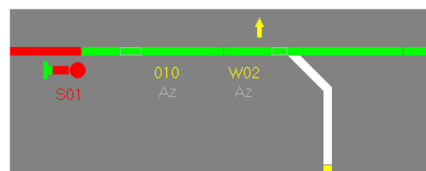
Dem FdI ist es nicht möglich seine Anlage selbstständig zu entstören. Um jedoch den Betrieb aufrechtzuerhalten muss er Hilfsbedienungen durchführen. Für den Versuch ist es wichtig, dass Störungen direkt nach Erkennen bearbeitet werden.

Störungsbetrieb

- Gleisfreimeldestörungen (Gleis zeigt „belegt“ an, obwohl es frei ist)



- Signalstörungen (Ausfall der Grünlampe, Signal kommt nicht mehr auf Fahrt)



Störungsbetrieb

- Grundsätzlich besteht immer die Möglichkeit den Zuglenkbetrieb für den kompletten Bedienplatz bzw. für einzelne Signal aus- und einzuschalten
- im vorliegenden Fall fällt der Zuglenkbetrieb störungsbedingt aus



Der Umgang mit den Störungen wird im Handout „Störbetrieb“ erklärt.


Vielen Dank für Ihre Aufmerksamkeit

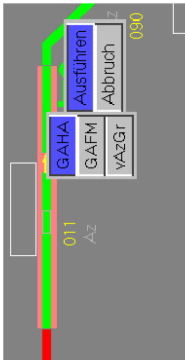

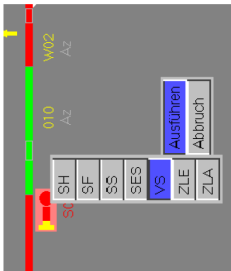



Fragen?

Anhang B: Handout zur Fehlerdiagnose (exemplarisch Originalbedienplatz)

Versuchsreihe Einführung Stellwerksbedienung

 ifev
Institut für Eisenbahnsignalung
und Verkehrssicherung
TU Duisburg-Essen

<p>Schritt 2 Auflösen Element Gleis</p> <ul style="list-style-type: none"> • Gleisabschnitt lfsauflösung (GAHA) • KF1+KF2 	
<p>Schritt 3 Fahrtstraße wie gewohnt einstellen</p> <ul style="list-style-type: none"> • Signal bleibt in „Halt“, jedoch mit einem gelben Fuß • Hilfsbedienebene ist ausreichend gesichert und aktiv 	
<p>Schritt 4 Vorsichtsignal stellen</p> <ul style="list-style-type: none"> • Abschnitt vor dem Signal <u>muss</u> mit einem Zug belegt sein (notwendige Bedingung) • Vorsichtsignal (VS) am Startsignal • KF1+KF2 	
<p>Schritt 5 Vorsichtsignal leuchtet</p> <ul style="list-style-type: none"> • VS leuchtet für eine begrenzte Zeit • Zug fährt vorbei • für weitere Zugfahrten weiter mit Schritt 1 bzw. Schritt 2 	

Ersteller: David Matz
Störungsbetrieb

15

Seite 2 von 5

Versuchsreihe Einführung Stellwerksbedienung

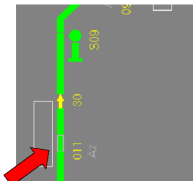
ifEV
Institut für Eisenbahnen
und Verkehrsicherung
TU Braunschweig

1. Gleisfreimeldestörung (Fahrwegelement gestört)

SICAS (Elektronisches Stellwerk)

Ausgangssituation

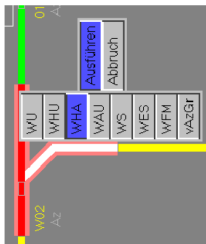
- Gleisfreimeldung der Weiche 2 (W02) ist nach Zugfahrt gestört
- Verschlussmelder der Elemente Weiche 2 (W02) und Gleis 11 (011) noch aktiv (Fahrstraße hat nicht aufgelöst)



Schritt 1

Auflösen Element Weiche
(nur wenn Weiche im Fahrweg verschlossen ist)


- Weichen Hilfsauflösung (WHA)
- Kommandofreigabe (Erklärung siehe nächstes Zeile)



Zur Verfahrenssicherung muss der FdL, bevor die Hilfsbedienhandlung durchgeführt werden kann, die Richtigkeit der Anzeige und des Bedienkommandos überprüfen. Hierfür erfolgt, vorab, eine Kommandofreigabe.

Der Bediener hat für die KF1- und KF2-Bedienelemente jeweils 15 s Zeit. Erfolgt keine Eingabe, wird der Befehl abgebrochen. Die Anzeige der Abfrage erfolgt im linken unteren Bildbereich.

Befehl und Anzeige prüfen => KF1 betätigen => nochmals prüfen => KF 2 betätigen



Ersteller: David Matz
Störungsbetrieb

2015

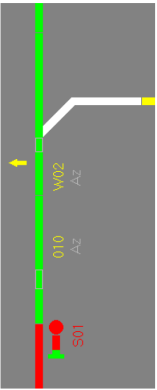
Seite 1 von 5

2. Signallöschung (Grünlampe defekt)

SICAS (Elektronisches Stellwerk)

Ausgangssituation

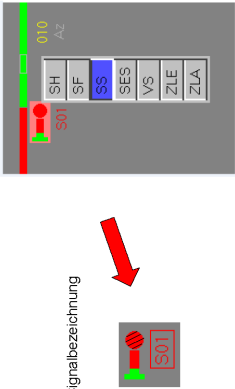
- Fahnsstraße ist eingelaufen (Fahweg grün ausgeleuchtet), aber Signal geht nicht auf Fahrt
- Signal bleibt in „Halt“, jedoch mit einem grünen Fuß und einem blinkenden Signalkopf
- Fahnsstraßenebene ist ausreichend gesichert und aktiv



Schritt 1

Signalsperren

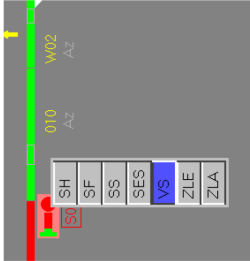
- Signal sperren (SS)
- roter Kasten erscheint um die Signalbezeichnung



Schritt 2

Vorsichtssignal stellen

- Abschnitt vor dem Signal muss mit einem Zug belegt sein (notwendige Bedingung)
- Vorsichtssignal (VS) am Startsignal
- KF1+KF2



Ersteller: David Maiz

Störungsbetrieb

2015

Seite 3 von 5

Versuchsreihe

Einführung Stellwerksbedienung

ifef

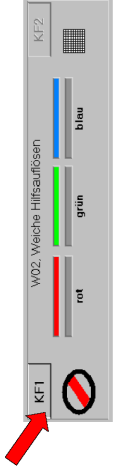
Institut für Eisenbahnen
und Verkehrssicherung

TU Braunschweig

Zur Verfahrenssicherung muss der FdI, bevor die Hilfsbedienhandlung durchgeführt werden kann, die Richtigkeit der Anzeige und des Bedienkommands überprüfen. Hierfür erfolgt, vorab, eine Kommandofreigabe.

Der Bediener hat für die KF1- und KF2-Bedienung jeweils 15 s Zeit. Erfolgt keine Eingabe, wird der Befehl abgebrochen. Die Anzeige der Abfrage erfolgt im linken unteren Bildbereich.


Befehl und Anzeige prüfen => KF1 betätigen => nochmals prüfen => KF2 betätigen



Schritt 3

Vorsichtssignal leuchtet

- VS leuchtet für eine begrenzte Zeit
- Zug fährt vorbei
- für weitere Zugfahrten weiter mit Schritt 2

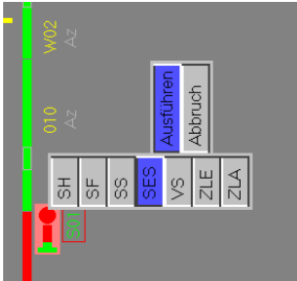


Schritt 4

Signal entsperren

(wenn Störung beseitigt wurde)

- Signal entsperren (SES)
- KF1+KF2



Ersteller: David Maiz

Störungsbetrieb

2015

Seite 4 von 5


3. Zuglenkbetrieb einzelner Signale ausgefallen


SICAS (Elektronisches Stellwerk)

Ausgangssituation

- Grundsätzlich besteht immer die Möglichkeit den Zuglenkbetrieb für den kompletten Bedienplatz bzw. für einzelne Signal aus- und einzuschalten
- hierbei wechselt die Signalbezeichnung von grün nach rot
- im vorliegenden Fall fällt der Zuglenkbetrieb störungsbedingt aus

S27



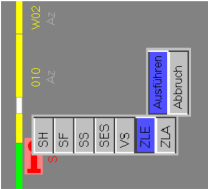


S27

Schritt 1

Zuglenkbetrieb einschalten

- Zuglenkbetrieb einschalten (ZLE)
- Signalbezeichnung wechselt von rot nach grün



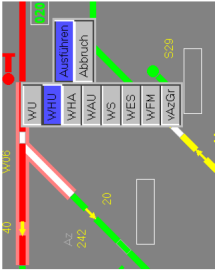
4. weitere Hilfsbehandlungen

SICAS (Elektronisches Stellwerk)

gestörte Weiche umstellen

(Ist die Gleisfreimeldung einer Weiche gestört, kann der FdI die Weiche mittels Hilfsbedienhandlung in die andere Lage bringen.)

- Ist die Weiche noch im Verschluss (Verschlussmelder leuchtet) ist dieser vorher zurückzunehmen (siehe Schritt Auflösen Element Weiche)
- Weiche Hilfsumstellen (WHU)
- KF1+KF2
- solange die Gleisfreimeldung der Weiche gestört ist, muss die Weiche mit der WHU umgestellt werden



Anhang C1: User Experience Questionnaire (UEQ)

Bitte geben Sie nun Ihre Einschätzung des Produkts ab. Kreuzen Sie bitte nur einen Kreis pro Zeile an.

	1	2	3	4	5	6	7
unbefriedigend	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
unverständlich	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
kreativ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
leicht zu lernen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
wertvoll	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
langweilig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
uninteressant	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
unberechenbar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
schnell	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
originell	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
behindernd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
kompliziert	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
abstoßend	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
herkömmlich	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
unangenehm	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
sicher	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
aktivierend	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
erwartungskonform	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ineffizient	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
übersichtlich	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
unpragmatisch	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
aufgeräumt	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
attraktiv	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
sympathisch	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
konservativ	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

UEQ_german.doc

Bitte geben Sie Ihre Beurteilung ab.

Um das Produkt zu bewerten, füllen Sie bitte den nachfolgenden Fragebogen aus. Er besteht aus Gegensatzpaaren von Eigenschaften, die das Produkt haben kann. Abstufungen zwischen den Gegensätzen sind durch Kreise dargestellt. Durch Ankreuzen eines dieser Kreise können Sie Ihre Zustimmung zu einem Begriff äußern.

Beispiel:

attraktiv	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	unattraktiv
-----------	-----------------------	----------------------------------	-----------------------	-----------------------	-----------------------	-----------------------	-------------

Mit dieser Beurteilung sagen Sie aus, dass Sie das Produkt eher attraktiv als unattraktiv einschätzen.

Entscheiden Sie möglichst spontan. Es ist wichtig, dass Sie nicht lange über die Begriffe nachdenken, damit Ihre unmittelbare Einschätzung zum Tragen kommt.

Bitte kreuzen Sie immer eine Antwort an, auch wenn Sie bei der Einschätzung zu einem Begriffspaar unsicher sind oder finden, dass es nicht so gut zum Produkt passt.

Es gibt keine „richtige“ oder „falsche“ Antwort. Ihre persönliche Meinung zählt!

Anhang C2: Fragebogen zur Beurteilung des Trainings

6. Wie sehr unterstützt die farbliche Gestaltung bei der Erkennung von Störungen?

sehr wenig		wenig		mittel		stark		sehr stark						
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

7. Wie sehr mussten Sie sich anstrengen, um die Bedienung zu erlernen?

sehr wenig		wenig		mittel		stark		sehr stark						
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

8. Was war für Sie besonders anstrengend während der Bedienung?

9. Was hat Sie an der Darstellungsart am meisten gestört?

Erstellt: David Metz

Seite 2

Fragebogen

zur Beurteilung des Trainings

VP-Nr.: _____

ifEV
Institut für Erwachsenen-
und Weiterbildung
TU Braunschweig

Im folgenden Fragebogen finden Sie Fragen, die sich auf unterschiedliche Aspekte während des Trainings beziehen.

Die abgebildeten Skalen sind in zwei Stufen zu verstehen. Wählen Sie bitte bei jeder Frage zunächst eine der Hauptkategorien und kreuzen Sie dann eine der drei dazugehörigen Zahlen an, die am ehesten beschreibt, wie Sie das Training fanden.

Für unsere Auswertung ist es sehr wichtig, dass Sie ehrlich antworten und keine Frage auslassen.

Beanspruchung während des Trainings

1. Wie anstrengend war das Erlernen der Software insgesamt?

sehr wenig			wenig			mittel			anstrengend			sehr anstrengend		
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

2. Wie hoch waren die geistigen Anforderungen während des Erlernens?

sehr niedrig			niedrig			mittel			hoch			sehr hoch		
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

3. Wie hoch waren die körperlichen Anforderungen während des Erlernens?

sehr niedrig			niedrig			mittel			hoch			sehr hoch		
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

4. Wie angenehm ist die farbliche Gestaltung der Oberfläche insgesamt?

sehr unangenehm			unangenehm			mittel			angenehm			sehr angenehm		
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

5. Wie angenehm ist der Kontrast der Oberfläche?

sehr unangenehm			unangenehm			mittel			angenehm			sehr angenehm		
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Erstellt: David Mitz

Seite 1

Anhang C3: Mehrdimensionaler Befindlichkeitsfragebogen (Kurzform A)

Bitte kreuzen Sie spontan ohne lange Überlegung in jeder Zeile jeweils einen Kreis an, der am ehesten auf Ihr momentanes Befinden zutrifft.

Im Moment fühle ich mich...	überhaupt nicht	1	2	3	4	sehr
zufrieden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ausgeruht	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ruhelos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
schlecht	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
schlapp	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
gelassen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
müde	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
unruhig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
munter	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
unwohl	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
entspannt	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Kurzform A

VP-Nr.

Datum

Geschlecht

w

O

m

O

Alter

Instruktion

Im Folgenden finden Sie eine Liste von Wörtern, die verschiedene Stimmungen beschreiben.

Bitte gehen Sie die Wörter der Liste nacheinander durch und kreuzen Sie bei jedem Wort das Kästchen an, das die augenblickliche Stärke Ihrer Stimmung am besten beschreibt.

Bitte beachten Sie:

- In der Liste sind mehrere Adjektive enthalten, die möglicherweise dieselbe oder eine ähnliche Stimmung beschreiben. Lassen Sie sich dadurch nicht verwirren, und geben Sie Ihre Antwort bei jedem Adjektiv unabhängig davon, wie Sie bei einem anderen Adjektiv geantwortet haben.
- Beurteilen Sie nur, wie Sie sich augenblicklich fühlen, nicht wie Sie sich im Allgemeinen oder gelegentlich fühlen.
- Wenn Ihnen die Antwort schwerfallen sollte, geben Sie die Antwort ab, die am ehesten zutrifft. Lassen sie keins der Wörter aus!

Anhang C4: Mehrdimensionaler Befindlichkeitsfragebogen (Kurzform B)


VP-Nr.

Datum

Geschlecht

w O

m O



Instruktion

Im Folgenden finden Sie eine Liste von Wörtern, die verschiedene Stimmungen beschreiben.
Bitte gehen Sie die Wörter der Liste nacheinander durch und kreuzen Sie bei jedem Wort das Kästchen an, das die augenblickliche Stärke Ihrer Stimmung am besten beschreibt.

Bitte beachten Sie:

- In der Liste sind mehrere Adjektive enthalten, die möglicherweise dieselbe oder eine ähnliche Stimmung beschreiben. Lassen Sie sich dadurch nicht verwirren, und geben Sie Ihre Antwort bei jedem Adjektiv unabhängig davon, wie Sie bei einem anderen Adjektiv geantwortet haben.
- Beurteilen Sie nur, wie Sie sich augenblicklich fühlen, nicht wie Sie sich im Allgemeinen oder gelegentlich fühlen.
- Wenn Ihnen die Antwort schwerfallen sollte, geben Sie die Antwort ab, die am ehesten zutrifft. Lassen sie keins der Wörter aus!

Bitte kreuzen Sie spontan ohne lange Überlegung in jeder Zeile jeweils einen Kreis an, der am ehesten auf Ihr momentanes Befinden zutrifft.

Im Moment fühle ich mich...	überhaupt nicht	1	2	3	4	sehr
schläfrig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
wohl	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ausgeglichen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
unglücklich	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
wach	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
unzufrieden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
angespannt	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
frisch	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
glücklich	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
nervös	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ermattet	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ruhig	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Kurzform B

Anhang C5: Fragebogen zur Beurteilung der Bedienungssicherheit

Die abgebildeten Skalen sind in zwei Stufen zu verstehen. **Wählen Sie bitte bei jeder Frage zunächst eine der Hauptkategorien und kreuzen Sie dann eine der drei dazugehörigen Zahlen an, die am ehesten beschreibt, wie Sie das Training fanden.**

Wie sicher fühlen Sie sich aktuell mit der Bedienung der Simulation?

sehr unsicher			unsicher			mittel			sicher			sehr sicher		
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Anhang D: Verteilung der Unregelmäßigkeiten

Verteilung für den standardisierten Testlauf (Gesamtversuchsdauer 13 min)

Nr.	Element	Art der Störung	Eintrittszeit	Anmerkung
1	S08	Zuglenkbetrieb ausgefallen	00:54	
2	S23	Signalstörung	02:00	Störungseintritt
3	W02	Gleisfreimelde störung	02:25	Störungseintritt
4	S24	Signalstörung	03:00	Störungseintritt
5	S01	Zuglenkbetrieb ausgefallen	04:20	
	S23	Signalstörung	05:30	Störung behoben
	S24	Signalstörung	06:29	Störung behoben
	W02	Gleisfreimelde störung	06:48	Störung behoben
6	S25	Zuglenkbetrieb ausgefallen	07:10	
7	W04	Gleisfreimelde störung	08:54	Störungseintritt
8	S11	Zuglenkbetrieb ausgefallen	10:00	
	W04	Gleisfreimelde störung	12:30	Störung behoben

Verteilung für den Hauptversuch (Gesamtversuchsdauer 30 min)

Nr.	Element	Art der Störung	Eintrittszeit	Anmerkung
1	S25	Zuglenkbetrieb ausgefallen	02:15:00	
2	S01	Zuglenkbetrieb ausgefallen	02:32:00	
3	S24	Zuglenkbetrieb ausgefallen	03:11:00	
4	S01	Signalstörung	03:53:00	Störungseintritt
5	S25	Zuglenkbetrieb ausgefallen	04:20:00	
6	S24	Zuglenkbetrieb ausgefallen	04:43:00	
	S01	Signalstörung	05:40:00	Störung behoben
7	W04	Gleisfreimelde störung	06:15:00	Störungseintritt
8	S24	Zuglenkbetrieb ausgefallen	07:00:00	
9	S23	Zuglenkbetrieb ausgefallen	07:20:00	
10	S24	Signalstörung	08:00:00	Störungseintritt
11	S23	Signalstörung	09:35:00	Störungseintritt
12	S01	Zuglenkbetrieb ausgefallen	09:50:00	
	S24	Signalstörung	10:45:00	Störung behoben
	W04	Gleisfreimelde störung	12:20:00	Störung behoben
13	W06	Gleisfreimelde störung	12:55:00	Störungseintritt
	S23	Signalstörung	13:20:00	Störung behoben
14	S01	Zuglenkbetrieb ausgefallen	14:00:00	
15	S22	Zuglenkbetrieb ausgefallen	14:42:00	
16	S23	Zuglenkbetrieb ausgefallen	15:40:00	
17	W02	Gleisfreimelde störung	16:55:00	Störungseintritt
18	S23	Zuglenkbetrieb ausgefallen	17:33:00	
19	S22	Zuglenkbetrieb ausgefallen	17:53:00	
	W02	Gleisfreimelde störung	19:30:00	Störung behoben
20	S23	Zuglenkbetrieb ausgefallen	20:18:00	
	W06	Gleisfreimelde störung	22:00:00	Störung behoben
21	S23	Zuglenkbetrieb ausgefallen	23:08:00	
22	S01	Zuglenkbetrieb ausgefallen	23:33:00	
23	S22	Zuglenkbetrieb ausgefallen	23:45:00	
24	S25	Signalstörung	24:48:00	Störungseintritt
25	S01	Zuglenkbetrieb ausgefallen	25:59:00	
26	S08	Zuglenkbetrieb ausgefallen	26:15:00	
27	S23	Zuglenkbetrieb ausgefallen	27:50:00	
	S25	Signalstörung	28:00:00	Störung behoben
28	S01	Zuglenkbetrieb ausgefallen	28:20:00	

Anhang E: Tool zur Erstellung von Heatmaps

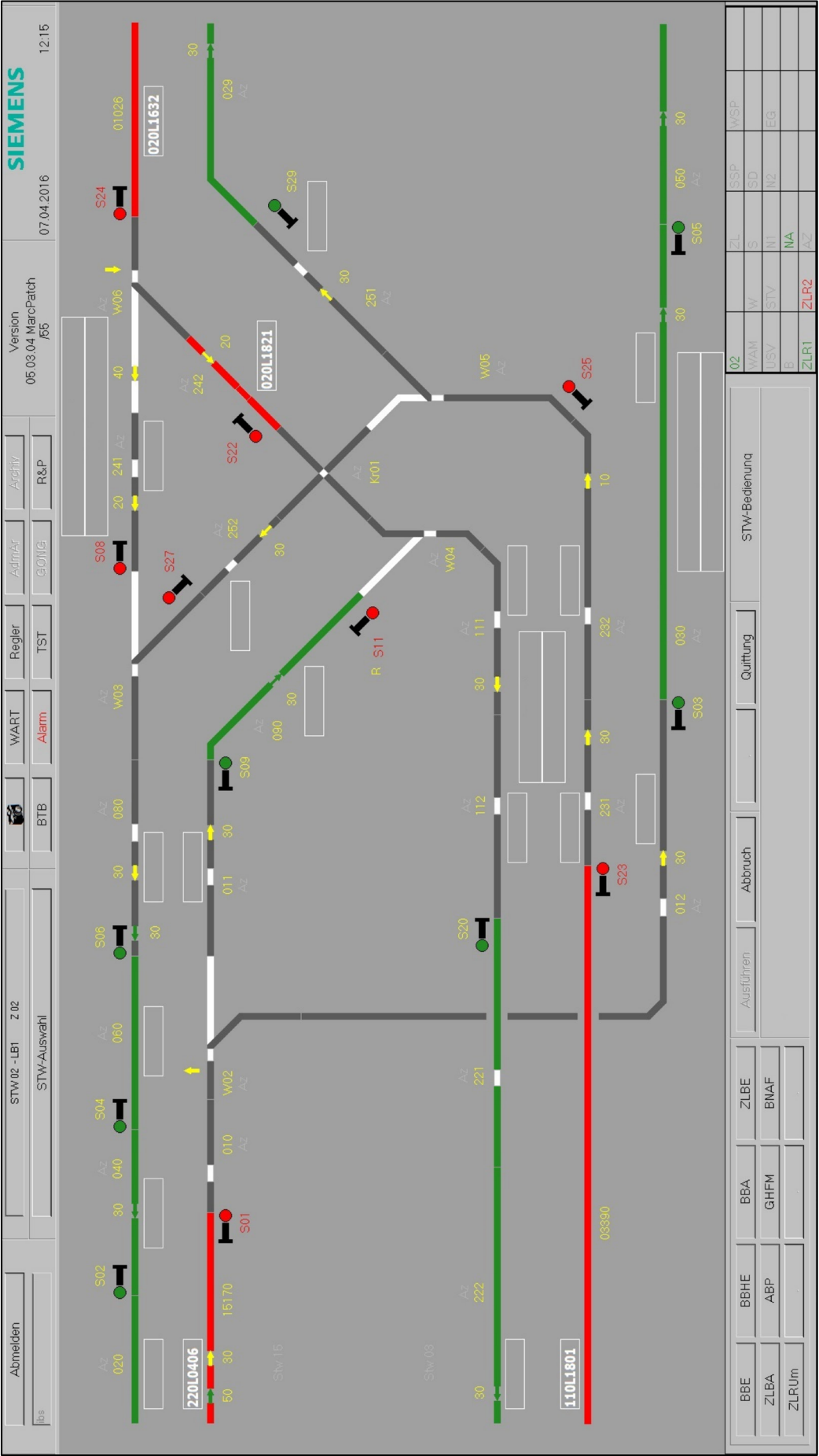
SIEMENS

Siemens Rail Automation
Version 0.6

Upload Custom Data

Upload Custom Picture

Upload your My Gaze Studie data with the "Upload Custom Data" button.
Use the "Upload Custom Picture" button to upload the picture of the application you tested with My Gaze



Designbedienplatz

Anhang G1: Ergebnisse User Experience

Tabelle A-1: Auswertung Fragebogen zur Beurteilung der User Experience (Test auf Normalverteilung und deskriptive Statistik)

Dimension	Test auf Normalverteilung			deskriptive Statistik (farbliche Gestaltung)			deskriptive Statistik (Geschlecht)			
	Shapiro-Wilk			Hell (N = 29)			Dunkel (N = 23)		männlich (N = 27)	
	Statistik	df	p-Wert	M	SD		M	SD	M	SD
Attraktivität	.955	52	.046	0.78	0.74	0.93	1.02	0.66	0.66	0.84
Durchschaubarkeit	.925	52	.003	0.85	1.10	1.23	1.26	1.00	0.76	1.06
Effizienz	.973	52	.287	0.84	0.81	1.16	1.15	0.86	0.81	0.78
Steuerbarkeit	.967	52	.153	1.16	0.87	1.01	1.38	0.79	0.78	0.85
Stimulation	.953	52	.039	1.01	0.92	0.98	1.11	0.72	0.87	1.01
Originalität	.975	52	.345	-0.02	0.94	0.30	-0.03	0.95	0.29	0.84

Tabelle A-2: Auswertung Fragebogen zur Beurteilung der User Experience (t-Test für unabhängige Stichproben)

Dimension	Levene-Test			t-Test (farbliche Gestaltung)			Levene-Test			t-Test (Geschlecht)		
	F			t			F			t		
		p-Wert	df		p-Wert	df		p-Wert	df		p-Wert	df
Attraktivität	0.012	.914	50	-0.739	.463	50	2.024	.161	50	-1.722	.091	50
Durchschaubarkeit	1.448	.235	50	-1.287	.204	50	0.047	.830	50	-1.747	.087	50
Effizienz	0.021	.886	50	-1.385	.172	50	0.030	.862	50	-1.484	.144	50
Steuerbarkeit	0.008	.930	50	0.579	.565	50	0.253	.617	50	-2.620	.012	50
Stimulation	0.000	.988	50	0.124	.902	50	0.398	.531	50	-.999	.323	50
Originalität	0.259	.613	50	-1.282	.206	50	0.355	.554	50	1.274	.209	50

Tabelle A-3: Auswertung Fragebogen zur Beurteilung der User Experience (U-Test)

Dimension	U-Test (farbliche Gestaltung)	U-Test (Geschlecht)
Attraktivität	$U(29, 23) = 306.5, p = .617; z = -0.500$	$U(27, 25) = 255, p = .129; z = -1.520$
Durchschaubarkeit	$U(29, 23) = 274.5, p = .273; z = -1.096$	$U(27, 25) = 250.5, p = .108; z = -1.607$
Effizienz	$U(29, 23) = 269, p = .232; z = -1.196$	$U(27, 25) = 266, p = .187; z = -1.318$
Steuerbarkeit	$U(29, 23) = 288.5, p = .404; z = -0.834$	$U(27, 25) = 210.5, p = .019; z = -2.340$
Stimulation	$U(29, 23) = 320, p = .802; z = -0.250$	$U(27, 25) = 302.5, p = .519; z = -0.645$
Originalität	$U(29, 23) = 272.5, p = .258; z = -1.131$	$U(27, 25) = 278.5, p = .277; z = -1.087$

Tabelle A-4: Auswertung Fragebogen zur Beurteilung der User Experience (Störungsbewertungen – Tests der Zwischensubjekteffekte)

Dimension	Multivariat (farbliche Gestaltung* Geschlecht)				
	df	Mittel der Quadrate	F	p-Wert	η^2
Attraktivität	1	1.370	2.494	.121	.049
Durchschaubarkeit	1	0.661	0.632	.430	.013
Effizienz	1	0.239	0.359	.552	.007
Steuerbarkeit	1	1.205	1.790	.187	.036
Stimulation	1	0.387	0.496	.485	.010
Originalität	1	0.380	0.470	.496	.010

Anhang G2: Ergebnisse psychische Beanspruchung

Tabelle A-5: Auswertung Fragebogen zur Beurteilung der psychischen Beanspruchung (Test auf Normalverteilung und deskriptive Statistik)

Item	Test auf Normalverteilung			deskriptive Statistik (farbliche Gestaltung)			deskriptive Statistik (Geschlecht)					
	Shapiro-Wilk			Hell (N = 27)			Dunkel (N = 23)			männlich (N = 27)		
	Statistik	df	p-Wert	M	SD		M	SD		M	SD	
Item 1	.959	50	.078	6.70	2.43		6.70	2.84		5.67	2.59	7.91
Item 2	.978	50	.479	7.81	2.24		7.17	2.55		6.85	2.46	8.30
Item 3	.811	50	.000	2.81	1.96		3.09	2.21		2.70	2.48	3.22
Item 4	.926	50	.004	8.96	2.59		10.00	2.70		10.33	2.29	8.39
Item 5	.916	50	.002	9.48	2.55		11.43	2.11		10.59	2.61	10.13
Item 6	.927	50	.004	9.56	2.99		10.39	1.90		10.44	2.34	9.35
Item 7	.956	50	.060	8.41	2.41		7.00	2.30		7.11	2.56	8.52

Tabelle A-6: Auswertung Fragebogen zur Beurteilung der psychischen Beanspruchung (t-Test für unabhängige Stichproben)

Item	Levene-Test			t-Test (farbliche Gestaltung)			Levene-Test			t-Test (Geschlecht)		
	F	p-Wert		t	df	p-Wert	F			t	df	p-Wert
Item 1	1.335	.254		0.011	48	.991	1.717			3.351	48	.001
Item 2	0.718	.401		0.946	48	.349	1.296			2.233	48	.030
Item 3	0.010	.919		-0.461	48	.647	1.825			0.875	48	.386
Item 4	0.015	.905		-1.383	48	.173	1.113			-2.731	48	.009
Item 5	2.753	.104		-2.922	48	.005	0.165			-0.640	48	.525
Item 6	1.912	.173		-1.155	48	.254	0.050			-1.531	48	.132
Item 7	0.340	.562		2.105	48	.041	2.069			2.110	48	.040

Tabelle A-7: Auswertung Fragebogen zur Beurteilung der psychischen Beanspruchung (U-Test)

Nr.	Item	U-Test (farbliche Gestaltung)	U-Test (Geschlecht)
1	Wie anstrengend war das Erlernen der Software insgesamt?	$U(27, 23) = 300, p = .837; z = -0.206$	$U(27, 23) = 155, p = .002; z = -3.050$
2	Wie hoch waren die geistigen Anforderungen während des Erlernens	$U(27, 23) = 284, p = .603; z = -0.520$	$U(27, 23) = 211.5, p = .052; z = -1.944$
3	Wie hoch waren die körperlichen Anforderungen während des Erlernens?	$U(27, 23) = 286, p = .625; z = -0.488$	$U(27, 23) = 207, p = .039; z = -2.063$
4	Wie angenehm ist die farbliche Gestaltung der Oberfläche insgesamt?	$U(27, 23) = 223, p = .085; z = -1.722$	$U(27, 23) = 176.5, p = .008; z = -2.637$
5	Wie angenehm ist der Kontrast der Oberfläche?	$U(27, 23) = 177.5, p = .009; z = -2.628$	$U(27, 23) = 264, p = .358; z = -0.919$
6	Wie sehr unterstützt die farbliche Gestaltung bei der Erkennung von Störungen?	$U(27, 23) = 281.5, p = .566; z = -0.574$	$U(27, 23) = 262.5, p = .342; z = -0.949$
7	Wie sehr mussten Sie sich anstrengen, um die Bedienung zu erlernen?	$U(27, 23) = 199, p = .028; z = -2.199$	$U(27, 23) = 220.5, p = .076; z = -1.775$

Tabelle A-8: Auswertung Fragebogen zur Beurteilung der psychischen Beanspruchung (Störungsbewertungen - Tests der Zwischensubjekteffekte)

Item	Multivariat (farbliche Gestaltung*Geschlecht)					
	df	Mittel der Quadrate	F	p-Wert	η^2	Trennschärfe
Item 1	1	2.187	0.380	.541	.008	.093
Item 2	1	0.355	0.067	.797	.001	.057
Item 3	1	0.237	0.053	.819	.001	.056
Item 4	1	26.641	4.808	.033	.095	.574
Item 5	1	12.722	2.362	.131	.049	.325
Item 6	1	5.882	0.938	.338	.020	.158
Item 7	1	4.885	0.975	.329	.021	.162

Anhang G3: Ergebnisse Befindlichkeit

Tabelle A-9: Auswertung Fragebogen zur Beurteilung der Befindlichkeit (T-Test auf Normalverteilung und deskriptive Statistik)

Dimension	Test auf Normalverteilung			deskriptive Statistik (farbliche Gestaltung)				deskriptive Statistik (Geschlecht)			
	Shapiro-Wilk		p-Wert	Hell (N = 29)		Dunkel (N = 23)		männlich (N = 27)		weiblich (N = 25)	
	Statistik	df		M	SD	M	SD	M	SD	M	SD
GS vorher	.219	52	< .001	17.66	1.56	16.96	2.03	17.56	1.36	17.12	2.19
GS nachher	.286	52	< .001	17.03	2.03	16.91	2.13	17.37	2.00	16.56	2.06
WM vorher	.150	52	.005	16.03	2.72	14.04	3.18	15.59	2.50	14.68	3.54
WM nachher	.131	52	.027	15.69	3.05	14.96	3.14	15.89	2.99	14.80	3.14
RU vorher	.171	52	.001	15.68	2.62	15.52	2.98	15.93	2.58	15.28	2.97
RU nachher	.135	52	.019	14.76	2.94	14.44	3.07	15.59	2.79	13.56	2.84
Differenz GS	.173	52	< .001	-0.62	2.11	-0.04	1.52	-0.19	1.92	-0.56	1.85
Differenz WM	.126	52	.039	-0.34	3.18	0.91	2.68	0.30	2.95	0.12	3.12
Differenz RU	.123	52	.049	-0.93	3.57	-1.09	2.41	-0.33	3.32	-1.72	2.69

Tabelle A-10: Auswertung Fragebogen zur Beurteilung der Befindlichkeit (t-Test für unabhängige Stichproben)

Dimension	Levene-Test		t-Test (farbliche Gestaltung)		Levene-Test		t-Test (Geschlecht)	
	F	p-Wert	t	df	F	p-Wert	t	df
GS vorher	0.788	.379	1.410	50	4.192	.046	0.858	39.22
GS nachher	0.000	.993	0.210	50	0.692	.409	-1.437	50
WM vorher	0.517	.475	2.432	50	3.611	.063	1.056	43.29
WM nachher	0.299	.587	0.850	50	0.322	.573	-1.281	50
RU vorher	0.207	.651	0.216	50	0.612	.438	0.841	50
RU nachher	0.053	.819	0.387	50	0.107	.745	-2.600	50
Differenz GS	3.364	.073	-1.103	50	0.003	.955	-0.715	50
Differenz WM	0.551	.461	-1.518	50	0.155	.695	-0.209	50
Differenz RU	2.072	.156	-0.180	50	0.919	.342	-1.649	50

Tabelle A-11: Auswertung Fragebogen zur Beurteilung der Befindlichkeit bezogen auf die farbliche Gestaltung (t-Test für abhängige Stichproben)

Dimension	Hell (N = 29)			Dunkel (N = 23)		
	M	SD	t	M	SD	t
GSv-GSn	-0.62	2.11	1.583	-0.04	1.52	0.137
WMv-WMn	-0.34	3.18	0.585	0.91	2.68	-1.635
RUV-RUn	-0.93	3.57	1.406	-1.09	2.41	2.163

Tabelle A-12: Auswertung Fragebogen zur Beurteilung der Befindlichkeit bezogen auf das Geschlecht (t-Test für abhängige Stichproben)

Dimension	männlich (N = 27)			weiblich (N = 25)		
	M	SD	t	M	SD	t
GSv-GSn	0.19	1.92	0.0501	0.56	1.85	1.513
WMv-WMn	-0.30	2.95	-0.523	-0.12	3.13	-0.192
RUV-RUn	0.33	3.32	0.522	1.72	2.69	3.203

Tabelle A-13: Auswertung Fragebogen zur Beurteilung der Befindlichkeit (U-Test)

Dimension	U-Test (farbliche Gestaltung)	U-Test (Geschlecht)
GS vorher	$U(29, 23) = 289.5, p = .402; z = -0.838$	$U(27, 25) = 326.5, p = .835; z = -0.208$
GS nachher	$U(29, 23) = 314.5, p = .717; z = -0.362$	$U(27, 25) = 251, p = .101; z = -1.640$
WM vorher	$U(29, 23) = 213, p = .025; z = -2.236$	$U(27, 25) = 293, p = .412; z = -0.821$
WM nachher	$U(29, 23) = 291, p = .431; z = -0.788$	$U(27, 25) = 268.5, p = .204; z = -1.271$
RU vorher	$U(29, 23) = 320.5, p = .809; z = -0.242$	$U(27, 25) = 290, p = .379; z = -0.879$
RU nachher	$U(29, 23) = 310.5, p = .670; z = -0.427$	$U(27, 25) = 197.5, p = .010; z = -2.581$
Differenz GS	$U(29, 23) = 240.5, p = .079; z = -1.755$	$U(27, 25) = 270.5, p = .209; z = -1.257$
Differenz WM	$U(29, 23) = 258, p = .161; z = -1.403$	$U(27, 25) = 328, p = .861; z = -0.175$
Differenz RU	$U(29, 23) = 300, p = .533; z = -0.623$	$U(27, 25) = 227, p = .041; z = -2.043$

Tabelle A-14: Auswertung Fragebogen zur Beurteilung der Befindlichkeit (Störungsbewertungen – Tests der Zwischensubjekteffekte)

Dimension	Multivariat (farbliche Gestaltung* Geschlecht)				
	df	Mittel der Quadrate	F	p-Wert	η^2
GS vorher	1	8.671	2.833	.099	.056
GS nachher	1	0.556	0.130	.720	.003
WM vorher	1	0.373	0.042	.838	.001
WM nachher	1	6.359	0.668	.418	.014
RU vorher	1	42.238	5.955	.018	.110
RU nachher	1	23.335	3.003	.090	.059
Differenz GS	1	4.835	1.376	.247	.028
Differenz WM	1	3.653	0.402	.529	.008
Differenz RU	1	2.784	0.293	.591	.006

Anhang G4: Ergebnisse Bedienungssicherheit

Tabelle A-15: Auswertung Fragebogen zur Beurteilung der Bedienungssicherheit (T-Test auf Normalverteilung und deskriptive Statistik)

Messpunkt	Test auf Normalverteilung			deskriptive Statistik (farbliche Gestaltung)			deskriptive Statistik (Geschlecht)					
	Shapiro-Wilk			Hell (N = 27)			Dunkel (N = 23)			männlich (N = 27)		
	Statistik	df	p-Wert	M	SD		M	SD		M	SD	weiblich (N = 23)
Messpunkt 1	.970	50	.232	9.33	2.77		9.22	2.37		10.30	2.30	8.09
Messpunkt 2	.946	50	.024	9.81	2.40		10.00	1.68		10.63	1.80	9.04
Messpunkt 3	.933	50	.007	10.78	1.74		10.96	1.94		11.48	1.53	10.13
Messpunkt 4	.941	50	.015	11.67	1.73		11.30	2.36		12.37	1.52	10.48
Differenz MP2/MP1	.915	50	.002	0.48	1.97		0.78	1.45		0.33	1.92	0.96
Differenz MP3/MP2	.702	50	.000	0.96	1.61		0.96	1.26		0.85	1.70	1.09
Differenz MP4/MP3	.698	50	.000	0.89	1.37		0.35	2.06		0.89	1.16	0.35
Differenz MP4/MP1	.975	50	.369	2.33	2.90		2.09	2.04		2.07	2.30	2.39

Tabelle A-16: Auswertung Fragebogen zur Beurteilung der Bedienungssicherheit (t-Test für unabhängige Stichproben)

Messpunkt	Levene-Test		t-Test (farbliche Gestaltung)			Levene-Test		t-Test (Geschlecht)		
	F	p-Wert	t	df	p-Wert	F	p-Wert	t	df	p-Wert
Messpunkt 1	0.336	.565	0.157	48	.876	0.006	.939	-3.323	48	.002
Messpunkt 2	3.788	.057	-0.310	48	.758	0.755	.389	-2.877	48	.006
Messpunkt 3	0.610	.439	-0.343	48	.733	1.590	.213	-2.794	48	.007
Messpunkt 4	2.202	.144	0.624	48	.535	2.520	.119	-3.675	48	.001
Differenz MP2/MP1	1.873	.177	-0.607	48	.547	1.255	.268	1.273	48	.209
Differenz MP3/MP2	1.170	.285	0.016	48	.988	1.799	.186	0.853	48	.398
Differenz MP4/MP3	0.702	.406	1.109	48	.273	0.132	.718	-1.080	48	.285
Differenz MP4/MP1	3.537	.066	0.342	48	.734	1.490	.228	0.440	48	.662

Tabelle A-17: Auswertung Fragebogen zur Beurteilung der Bedienungssicherheit (U-Test)

Messpunkt	U-Test (farbliche Gestaltung)	U-Test (Geschlecht)
Messpunkt 1	$U(27, 23) = 298, p = .806; z = -0.245$	$U(27, 23) = 160, p = .003; z = -2.950$
Messpunkt 2	$U(27, 23) = 299, p = .820; z = -0.228$	$U(27, 23) = 165, p = .004; z = -2.881$
Messpunkt 3	$U(27, 23) = 297, p = .789. z = -0.268$	$U(27, 23) = 184.5, p = .012; z = -2.499$
Messpunkt 4	$U(27, 23) = 292.5, p = .723; z = -0.355$	$U(27, 23) = 148.5, p = .001; z = -3.193$
Differenz MP2/MP1	$U(27, 23) = 293.5, p = .735; z = -0.338$	$U(27, 23) = 263, p = .345; z = -0.950$
Differenz MP3/MP2	$U(27, 23) = 282.5, p = .577; z = -0.558$	$U(27, 23) = 276.5, p = .498; z = -0.678$
Differenz MP4/MP3	$U(27, 23) = 286.5, p = .629; z = -0.483$	$U(27, 23) = 273, p = .451; z = -0.754$
Differenz MP4/MP1	$U(27, 23) = 303, p = .883; z = -0.148$	$U(27, 23) = 306, p = .929; z = -0.089$

Tabelle A-18: Auswertung Fragebogen zur Beurteilung der Bedienungssicherheit (Störungsbewertungen – Tests der Zwischensubjektffekte)

Messpunkt	Multivariat (farbliche Gestaltung*Geschlecht)					
	df	Mittel der Quadrate	F	p-Wert	η^2	Trennschärfe
Messpunkt 1	1	1.531	0.269	.606	.006	.080
Messpunkt 2	1	0.594	0.153	.698	.003	.067
Messpunkt 3	1	0.734	0.246	.622	.005	.077
Messpunkt 4	1	7.835	2.404	.128	.050	.330
Differenz MP2/MP1	1	0.218	0.070	.792	.002	.058
Differenz MP3/MP2	1	4.456	0.471	.496	.010	.050
Differenz MP4/MP3	1	32.275	2.894	.096	.059	.582
Differenz MP4/MP1	1	16.294	2.567	.116	.053	.348

Tabelle A-22: Auswertung objektiver Leistungskennzahlen (t-Test für unabhängige Stichproben)

Kennzahlen	Levene-Test		t-Test (farbliche Gestaltung)			Levene-Test		t-Test (Geschlecht)		
	F	p-Wert	t	df	p-Wert	F	p-Wert	t	df	p-Wert
abgefertigte Zugfahrten	2.476	.124	-1.690	38	.099	16.640	< .001	-1.903	19.452	.072
Gesamtfehler-entdeckungszeit	0.473	.496	0.281	38	.780	0.800	.377	-0.188	38	.852
KFV	2.097	.156	-2.213	37	.033	0.216	.163	0.871	37	.871
Anzahl der Hilfs-handlungen	2.097	.156	-1.507	38	.140	0.158	.693	-1.059	38	.296
Anzahl der Bedienfehler	1.699	.200	1.492	38	.144	0.898	.349	1.413	38	.166
Gesamt-überwachungszeit	5.910	.020	1.151	29.80	.259	0.225	.638	0.532	35	.598

Tabelle A-23: Auswertung objektiver Leistungskennzahlen (U-Test)

Kennzahlen	U-Test (farbliche Gestaltung)	U-Test (Geschlecht)
abgefertigte Zugfahrten	$U(21, 19) = 133, p = .074; z = -1.812$	$U(23, 17) = 160, p = .342; z = -0.977$
Gesamtfehler-entdeckungszeit	$U(21, 19) = 191, p = .830; z = -0.230$	$U(21, 19) = 173, p = .551; z = -0.616$
KFV	$U(20, 19) = 94, p = .006; z = -2.698$	$U(23, 16) = 159, p = .489; z = -0.714$
Anzahl der Hilfs-handlungen	$U(21, 19) = 152.5, p = .205; z = -1.283$	$U(23, 17) = 157.5, p = .302; z = -1.048$
Anzahl der Bedienfehler	$U(21, 19) = 151.5, p = .196; z = -1.307$	$U(23, 17) = 139, p = .126; z = -1.554$
Gesamtüberwachungszeit	$U(19, 18) = 125, p = .169; z = -1.398$	$U(22, 15) = 151, p = .680; z = -.433$

